UC San Diego
UC San Diego Electronic Theses and Dissertations

Title
Efficient Patient Specific Image Guidance in Liver Surgery

Permalink
https://escholarship.org/uc/item/3979f2bg

Author
Barrow, Michael Joseph

Publication Date
2020

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA SAN DIEGO

Efficient Data Driven Patient Specific Image Guidance in Liver Surgery

A dissertation submitted in partial satisfaction of the
requirements for the degree
Doctor of Philosophy

in

Computer Science

by

Michael Joseph Barrow

Committee in charge:

Professor Ryan Kastner, Chair
Professor Sonia Ramamoorthy, Co-Chair
Professor Chung Kuan Cheng
Professor Petr Krysl
Professor Nadir Weibel

2020

Copyright
Michael Joseph Barrow, 2020
All rights reserved.

The dissertation of Michael Joseph Barrow is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Co-Chair

Chair

University of California, San Diego

2020

DEDICATION

To Mum and Co.

EPIGRAPH

See with your own eyes,

or you will see nothing

—?

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to thank all Kaster Research Group members, my research collaborators, my PhD committee members, my professional mentors at Intel, and my professional mentors at Arm for their precious comments and feedback on my research. All discussions we had were invaluable and significant. I also gratefully acknowledge that my work was supported by; the ACM SIGHPC/Intel Computational & Data Science Fellowship, the CTRI Innovative Technology Pilot Grant, Cognex Corporation, Intel Corporation, and Nvidia Corporation.

Chapter 2, in full, is a reprint of the material as it appears in the 14th International Symposium on Visual Computing (ISVC), Barrow, Michael; Ho, Nelson; Althoff, Alric; Tueller, Peter, and Kastner, Ryan, 2019. There are small changes in format and phrasing as a chapter within this larger paper. The dissertation author was the primary investigator and author of this paper.

Chapter 3, in full, is a reprint of the material as it appears in the 28th International Conference on Field Programmable Logic and Applications (FPL), Barrow, Michael; Burns, Steven, and Kastner, Ryan, 2018. There are small changes in format and phrasing as a chapter within this larger paper. The dissertation author was the primary investigator and author of this paper.

Chapter 4, in full, is a reprint of submitted material as it may appear in Medical Image Analysis (MIA), Barrow, Michael; Chao, Alice; Fowler, Kathryn; He, Qizhi; Ramamoorthy, Sonia, and Kastner, Ryan, 2020. There are small changes in format and phrasing as a chapter within this larger paper. The dissertation author was the primary investigator and author of this paper.

Chapter 5, in full, is a reprint of submitted material as it appears in the third Black in AI Workshop (BAI), Barrow, Michael; He Qizhi, and Kastner, Ryan, 2019. There are changes in format and phrasing as a chapter within this larger paper. The dissertation author was the primary investigator and author of this paper.

VITA

2008	Bachelor of Engineering, in Applied Computing and Electronics, Bournemouth University, Poole, England
2011	Master of Science, in Electronic Engineering, Hong Kong University of Science and Technology, Hong Kong
2020	Doctor of Philosophy, in Computer Science, University of California San Diego, La Jolla, United States of America

PUBLICATIONS

Barrow, Michael; Chao, Alice; Romano, Christopher; Liu, Shanglei; "Force feedback as a surgical learning aid: A preliminary study on suture tying", *Society of American Gastrointestinal and Endoscopic Surgeons Annual Meeting (SAGES)*, (accepted).

Barrow, Michael; Chao, Alice; He, Qizhi; Fowler, Kathryn; Ramamoorthy, Sonia; Kastner, Ryan. "MR Elastography Method for Personalized Surgical Image Guidance", *Journal of Medical Image Analysis (MIA)*, (under review).

Barrow, Michael; He, Qizhi; Kastner, Ryan. "Data Driven Tissue Models for Surgical Image Guidance", *Black in AI Workshop*, 2019.

Barrow, Michael; Ho, Nelson; Althoff, Alric; Tueller, Peter; Kastner, Ryan. "Benchmarking Video with the Surgical Image Registration Generator (SIRGn) Baseline", *International Symposium on Visual Computing (ISVC)*, 320-331, 2019.

Barrow, Michael; Burns, Steven; Kastner, Ryan. "A FPGA Accelerator for Real-Time 3D Non-rigid Registration Using Tree Reweighted Message Passing and Dynamic Markov Random Field Generation", *Field-Programmable Logic and Applications (FPL)*, 335-3357, 2018.

Richmond, Dustin; Barrow, Michael; Kastner, Ryan. "Everyone's a Critic: A Tool for Exploring RISC-V Projects", *Field-Programmable Logic and Applications (FPL) 2018*, 260-2604, 2018.

Barrow, Michael; Liu, Shanglei; Ho, Nelson; Kastner, Ryan; Tapia, Eduardo; Yang, Xinyi; Chen, Brendon; Horgan, Santiago; Ramamoorthy, Sonia. "A ground truth 3D video data set for augmented reality robotic MIS algorithms", *Society of American Gastrointestinal and Endoscopic Surgeons Annual Meeting (SAGES)*, 80808, 2017.

Barrow, Michael; Thomas, Shelby; Alvarado, Christine. "ERSP: A Structured CS Research Program for Early-College Students", *Innovation and Technology in Computer Science Education (ITiCSE)*, 148-153, 2016.

ABSTRACT OF THE DISSERTATION

Efficient Data Driven Patient Specific Image Guidance in Liver Surgery

by

Michael Joseph Barrow

Doctor of Philosophy in Computer Science

University of California San Diego, 2020

Professor Ryan Kastner, Chair
Professor Sonia Ramamoorthy, Co-Chair

Liver cancer is a life threatening disease that must be treated urgently once detected. A partial hepatectomy to remove cancerous lesions has become the mainstay of treatment. A crucial part of the surgery is to identify where the tumors, vessels, and other important landmarks are located.

Surgeons require years of training and practice to perfect the art of image guided surgery. The liver is soft and readily changes shape during a procedure. This means the surgeon must mentally map landmarks between preoperative scans and the surgical view of a liver by modeling the rotation, scaling and distortion of the liver shape in their mind's eye. Only then can they

estimate where landmarks are relative to their tools.

Computer-aided Image Guidance (CAIG) fuses preoperative scans with intraoperative images to provide more detailed information about the surgical site. During surgery, CAIG can merge preoperative data directly into the surgeon's view as a visual overlay on top of the intraoperative video feed.

The goal is to display the precise location of vessels, ducts and tumors that are hidden beneath the surface shown by the camera.This type of CAIG allows a surgeon to avoid damaging important vessels while excising a tumor, reducing patient risk. Although accurate real-time CAIG is a valuable and important surgical tool, no system exists that is efficient enough to maintain clinically acceptable accuracy and frame rate.

This work explores the idea that a given CAIG system can be optimized for the clinical requirements of each surgical case. Optimization is done using patient specific preoperative data to tune algorithm and hardware configurations. The thesis statement is that:

"Preoperative and intraoperative image data can be used to instruct clinically significant efficiency optimizations in a Computer Aided Image Guidance pipeline."

We demonstrate the thesis statement by discussing data driven efficiency improvements we have contributed to the canonical CAIG pipeline. Specifically, our methods improve efficiency of *"non-rigid 3D registration"* and *"physical simulation"* which are online steps used to track surgical landmarks at video rate in the setting of liver surgery image guidance.

Taken together our contributions collated in the thesis build a suite of new analysis and implementation approaches that have clinically significant impact on the efficiency of CAIG pipelines.

Chapter 1

Introduction

1.1 Significance Of Accurate Image Guidance In Liver Surgery

The World Health Organization estimates that over one million people will die from liver cancer in 2030 [9]. Worldwide, it ranks fourth in terms of cancer-related death and sixth with respect to incident cases [10]; it has the fastest growth of incidence and the second highest mortality of all cancers in the United States [11]. It has a five-year survival rate of 18% making it the second most lethal cancer [12]. Liver cancer will surpass breast, prostate, and colorectal cancer to become the third leading cause of cancer related deaths by 2030 [13].

Surgical resection of metastatic disease improves survival for several primary malignancies, including colorectal cancers, neuroendocrine tumors, and gastric cancers [14, 15, 16, 17, 18]. Resections range from small wedge resection to major hepatectomies. The ultimate goal is to remove all tumors and leave behind as much liver as possible. Such parenchymal sparing approaches lead to reduced hepatic dysfunction/failure, morbidity, length of hospitalization, and mortality compared to major hepatectomies [19]. However, achieving a complete resection through this sparing approach can be difficult. Clinical investigations suggest that microscopic

(R1) or macroscopic (R2) resections with tumor at the surgical margins occur in 15-60% of liver resections; this is likely attributable to a combination of imprecise imaging and a disconnect between the preoperative images and the ability to accurately localize the lesions in the operating room [20]. ***Image guidance is the limiting factor for accurately localizing landmarks and performing more precise liver surgeries.***

1.2 Limitations Of Traditional Image Guided Liver Surgery

Tumors are often spread throughout the liver, they may be buried deep beneath the surface, and they can be bundled within a network of blood vessels and bile ducts. In the metastatic setting, resection often follows chemotherapy and what were once clearly visible tumors can become difficult to see, i.e., isoechoic, on intraoperative ultrasound. This makes it difficult for surgeons to intraoperatively locate tumors. Thus, surgeons consult preoperative images to resect the area of the lesion by relying on anatomical landmarks like the tumor's relationship to vasculature [21, 22, 23, 24, 25, 26, 27, 28, 29, 30]. Magnetic resonance imaging is a highly sensitive imaging modality with added benefit of functionally relevant information like proton density fat fraction (PDFF), R2* analysis, and elastography data to inform liver steatosis, iron and fibrosis, respectively [31, 32, 33, 34, 35, 36, 37]. Unfortunately, mapping landmarks from the preoperative 2D cross-sectional images to intraoperative 2D ultrasound images and the 3D surface of the liver is challenging even for an experienced surgeon. This requires the visuospatial synthesis of cross-sectional anatomic relationships with actual anatomic relationships, an extremely difficult but crucial skill.

The complex relationships between tumors, vasculature, and other important surgical landmarks change during surgery. To start, the preoperative scans are done with the patient in one position while the surgery has the patient in a different position and orientation [38, 39, 40], which can lead to spatial changes of landmarks from the preoperative plan [41]. Furthermore, the liver can undergo large topological changes during surgery [42]. For instance the entire right liver may be fully mobilized from the right retroperitoneum and inferior vena cava. In turn, it can be rotated to the left, leading to a contorted shape that allows the surgeon to access posterior tumors and create oncologically adequate resection margin(s). Following every significant mobilization, surgeons must perform a manual "image-to-actual" virtual or mental registration process. Surgeons rely heavily upon their experience to determine the relationships between anatomical landmarks in preoperative images and the surgical scene [25, 43]. They often resort to primitive techniques as aids in locating the tumors, e.g., measuring intraoperative distances using rulers and cauterizing "place markers" on the liver surface.

1.3 Enhancing Guidance Accuracy With Computer Vision

Traditional laparoscopic image guidance is hampered by the extreme variability of the livers appearance caused by its compliant biomechanical properties. The surgeon is forced to mentaly map preoperative MRI/CT scans to their view of the surgical scene to track landmarks such as tumors and vessels. This task is extremely labor intensive, requires years of training, and is error prone. The goal of CAIG is automate this mental mapping by fusing together preoperative scans and intraoperative images from a laparoscope. Real-time registration of the preoperative 3D data into an orientation consistent with the current surgical viewpoint provides valuable information about the position of important surgical landmarks, which enables more precise surgical procedures. However, this is challenging. Liver biomechanics are complex and preoperative surgical landmarks cannot be mapped to the laparoscope video using simple heuristic rules. ***Landmark registration must be physically based in order to accurately position image guidance cues***. This requires that the intraoperative images from a laparoscopic camera are processed to become input to a physical simulation. Several hard computer vision problems must be solved at video rates to achieve this. A CAIG system must efficiently solve these hard problems with frame rates above 10Hz and accuracy margins of less than 5 millimeters to be of practical use [44]. In this thesis, we explore the efficiency of physically based surgical image guidance in liver surgery. CAIG image processing can be broken down into a pipeline of steps that map to common computer vision problems. A canonical pipeline of the required steps is shown in fig. 1.1. The four main steps are as follows:

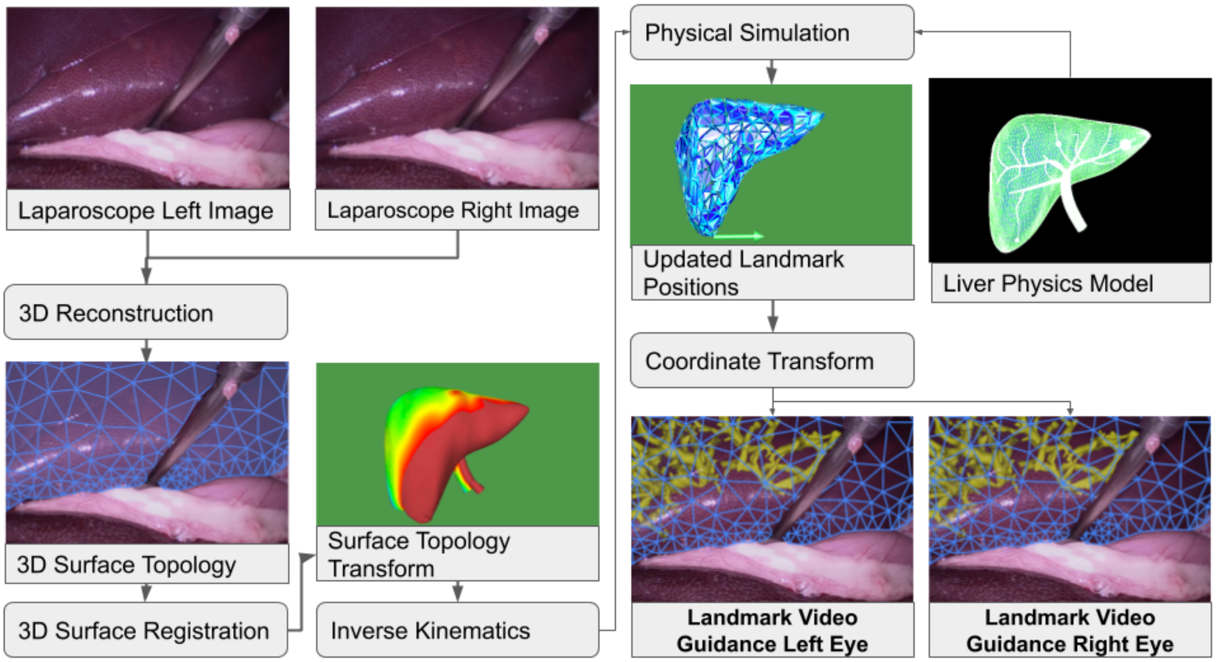


Figure 1.1: Canonical steps used to convert intraoperative video and preoperative MR scans into a dynamic video overlay of important surgical landmarks such as tumors and vessels. Step 1 **3D Reconstruction** converts traditional 2D video of the surgical scene to 3D so that the liver surface can be mapped to the volumetric preoperative liver image. Step 2 **3D Surface Registration** is the initial mapping of intraoperative to preoperative imaging where a non-rigid registration maps the coordinate system of the surgical liver images to the preoperative MR liver image. Step 3 **Physical Simulation** completes the mapping by using a physical simulation to project the unseen internal positions of the surgical landmarks the surgeon is interested in. The entire pipeline must run at video rate in order to give the surgeon Computer Vision assisted Image Guidance in real time. Step 4 **Display Projection** is the final step where surgical landmarks are rendered to a display after a coordinate space transform.

Step 1: The surface of the liver must be extracted before intraoperative images can be fused to preoperative 3D volumetric liver scans. A 3D reconstruction step reconstructs the surface of the liver from two dimensional RGB laparoscopic cameras. RGB cameras are the primary image sensor used in IG surgery, making this step canonical. Either a monocular (single camera) or stereoscopic (two camera) laparoscope may be used. Both require complicated algorithmic techniques to reconstruct a 3D coordinate system for use in later steps. Besides RGB, other wide field-of-view image sensors such as Infra Red, Laser, Structured Light, Time of Flight and hybrid devices are being explored for use in surgery [45, 46]. The variety of cameras, sensors and 3D reconstruction steps is a broad research area that is beyond the scope of this thesis. We assume a good quality 3D image is reconstructed in step 1 and focus on the following pipeline stages.

Step 2: The surface motion of the liver must be measured so as to track the motion of surgical landmarks and correctly position image guidance cues. A 3D surface registration step tracks surface motion of the liver. This step merges 3D scans of the surface from surgical video frames into the same coordinate system. Because the liver is soft and readily deforms, a "non-rigid" 3D surface registration is used to perform the mapping. This must also be a "global" registration. Every video frame from the laparoscope must be mapped to the same coordinate system so as to consistently position surgical landmarks in all video frames. A later step maps the preoperative liver images to the same coordinate space and allows the surgeon to locate landmarks of interest from those preoperative scans. The registration is also a real time step. The algorithm must finish in time to keep up with the 10Hz video frame rate required for practical image guidance.

Step 3: The position of landmarks beneath the visible surface such as vessels and tumors must be tracked to within an accuracy of 5 millimeters for CAIG to be clinically useful. Most of the liver is not visible in the surgical scene, meaning step 2 cannot perform a complete mapping between the surgeons view and preoperative scan. In order to accurately position surgical landmarks that are either out of the surgeons line of sight or contained within the organ, the effects of surface warping of the liver must be accurately projected throughout a virtual model containing the surgical landmarks of interest. A physical simulation of the liver uses the tracked surface movement to model internal movement of surgical landmarks so that they can be correctly positioned to the surgeon in a display.

Step 4: Once landmark positions have been calculated, they must be fused into the laparoscope video feed. The display projection step uses a camera transform matrix to translate from the physical simulator coordinate system to the laparoscope video frame. Next, the landmarks are visually presented to the surgeon. Although fig. 1.1 shows one method of displaying vessels, effective presentation methods for surgical guidance cues is an open research problem that is beyond the scope of this thesis [47, 48].

1.4 Thesis Statement

No physically based CAIG system exists that is efficient enough to maintain clinically acceptable accuracy and frame rate during the liver resections used to treat cancer. In general, it is challenging to implement CAIG that can meet clinical accuracy requirements at video frame rates. This is because the pipeline steps use computationally intensive algorithms to calculate accurate results. This creates two competing performance goals. On one hand, accuracy can be improved if the pipeline has more compute resource to refine the calculation during each step. On the other, frame rate can be improved by using less compute resource to refine calculations so that more steps can be completed in the same time period. Efficient CAIG pipelines that balance speed and accuracy in liver procedures are therefore an interesting open research problem.

This work explores the idea that a given CAIG system can be optimized for the clinical requirements of each patient. This is done using patient specific preoperative data to tune algorithm and hardware configuration. The thesis statement is that: ***Preoperative and intraoperative image data can be used to instruct clinically significant efficiency optimizations in a CAIG pipeline***.

1.5 Contemporary Physically Based Image Guidance

In this section we describe an example state of the art CAIG pipeline applied to telerobotic image guided liver surgery. During our discussion of the pipeline, we outline the specific research challenges that this thesis tackles and why they are important. Specifically we discuss the limitations of the state of the art in non-rigid registration and video rate liver simulations. These steps are used to track and position guidance landmarks respectively.

A great deal of the literature implements CAIG pipelines on a surgical robotics platform. Robots make experiments easy to repeat, and robots can be instrumented to record data more easily than a human surgeon. Furthermore, CAIG has a tangible benefit in robotic procedures. Robots can be controlled with sub millimeter precision, although this potential cannot be fully realized without accurate CAIG to guide them.

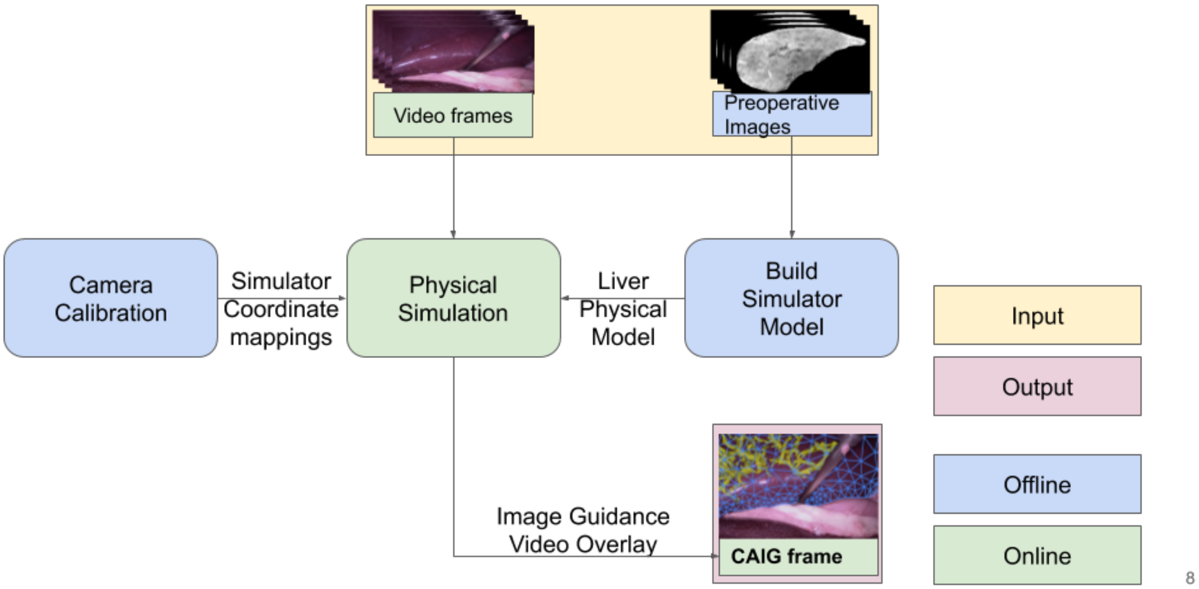


Figure 1.2: Basic block diagram of a telerobotic Computer Aided Image Guidance pipeline for use in liver procedures. System input is highlighted in yellow, system output is marked in red. Online and offline steps are green and blue respectively.

Figure 1.2 is a basic block diagram of a canonical state of the art CAIG system for telerobotic minimally invasive liver procedures. At a high level, preoperative images can be processed offline ahead of surgery to extract guidance cues. However, the intraoperative laparoscopic video must be processed online in real time as the surgery progresses. A physical simulation is used in an online step to fuse the intraoperative and preoperative images and is the key step in providing image guidance.

In one offline flow, preoperative images are used to model the biomechanics of the liver. These models are used by the simulator to position landmarks. In another offline flow, robot sensor data including camera intrinsic and extrinsic data is calibrated and fed into the online CAIG pipeline. This data is needed to translate between motions in the real world surgical scene and motions in the physical simulation domain.

Figure 1.3 shows a detailed flow of steps to implement a CAIG pipeline on a telerobotic surgical system. Blue steps are offline, or preoperative. Green steps are on-line and done in real time during surgery (intraoperative). We provide brief discussion of the steps in fig. 1.3 for background primer on the research contributions of this thesis.

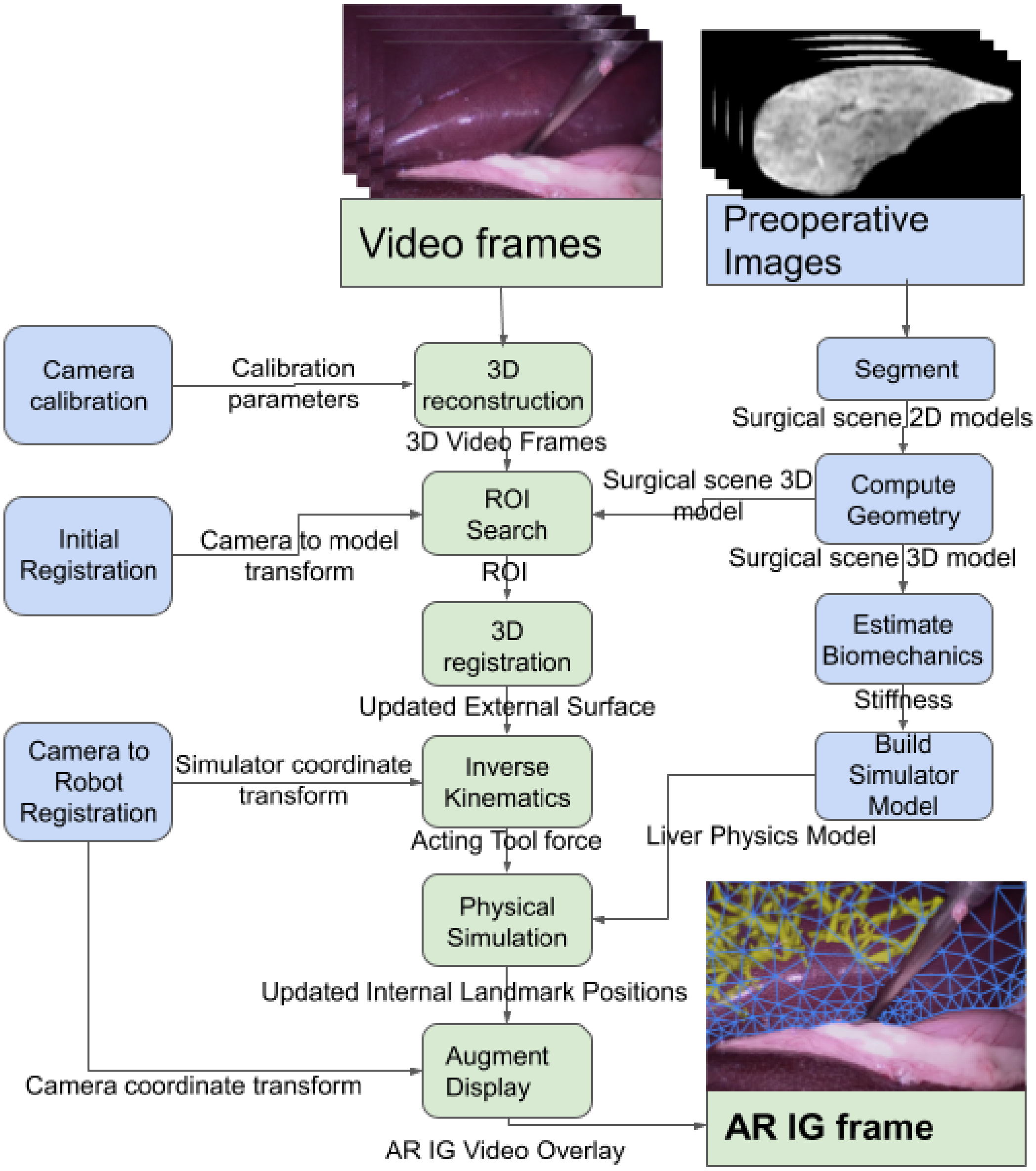


Figure 1.3: Detailed block diagram of a telerobotic computer aided image guidance pipeline. The intraoperative system input (yellow) is the stereo video feed of a surgical robot. Preoperative input (yellow) are MRI scans of the abdomen. Output is an augmented video stream with overlaid surgical landmarks (red). CAIG has both offline and online steps depicted in blue and green respectively. The system requires offline calibration to establish a mapping between the robot camera coordinate system and the physical simulation coordinate system. In an additional offline step, the surgical scene is extracted from preoperative images and used to create a physical model of the surgical scene. During surgery, online steps maintain registration between the scene model and video feed. Physical simulation is used in the online flow to propagate surface changes in the video feed within the liver so as to accurately position the CAIG landmark overlay

Segmentation

In this step, preoperative images are segmented into anatomical regions of interest. MRI is a common preoperative scan used in CAIG pipelines that is robust and accurate for metastic liver tissue imaging [49]. The liver structure must be segmented out from background information in the MRI before it can be used in the CAIG pipeline. A basic whole organ segmentation deliniates the gross anatomy or the exterior liver geometry. This step is sometime referred to as "volumetry" [50]. After the liver boundary is defined, a more granular segmentation is used to mark individual surgical landmarks, such as subcutaneous metastasis. During online steps, the CAIG pipeline tracks the position of these landmarks and presents them as guidance cues to the surgeon. Segmentation is routinely performed manually by imaging experts, but it can be automated or semi-automated. Automated segmentation of liver MRI is a popular research topic. Many methods have been proposed that can be categorized as Clustering, Statistical, Region Based, Energy Based, or Artificial Intelligence (AI) methods as defined by Kumar's taxonomy [51]. Much of this effort has been invested in tumor, lesion and vessel segmentation [52, 53, 54] and recent advances have been made in gross anatomy segmentation [55]. In general, good quality segmentation of MRI is routinely performed for traditional image guidance manually, with AI methods showing great potential.

Compute Geometry

After segmentation, the series of 2D MRI scans are assembled into volumetric 3D geometry. Volumetric geometry describes the spaces in between a series of MRI scans (typically 10 millimeters) so that the liver anatomy is fully defined in the 3D domain. In other words, continuous surfaces can "skin" the exterior of segmented MRI structures to create solid geometry. It is necessary to define solid geometry so that the CAIG pipeline can visually display landmarks at any orientation. The surfaces are also needed to construct liver models for the CAIG simulator.

This step is not always explicit. Some segmentation steps compute 3D geometry as a byproduct [56]. Other methods skip the segmentation and compute geometry directly from preoperative scans [57]. Regardless, the method chosen to compute geometry must represent the liver in a way compatible with the CAIG physical simulation step. Volumetric mesh based methods such as Finite Element Analysis FEA are common in physically based CAIG [58, 59]. The compute geometry step therefore typically generates mesh geometry, which can then be decomposed into basic elements for simulation [60]. Besides being naturally compatible with simulators, the mesh has other desirable properties. First, is simple which lends itself to video-rate processing. Second, it defines the surface of liver geometry explicitly, which is useful when computing collisions with other objects such as surgical tools during the simulation of liver procedures.

Common meshing approaches include 3D Delaunay triangulation [61], geodesic approaches [56], greedy strategies [62], and end to end image based methods [57]. Generating a high quality mesh is a well studied problem that is routinely performed [63].

Estimate Biomechanics

Biomechanics model the stress strain relationship of liver tissue. This model is used by the physical simulator to calculate the displacement of image guidance landmarks caused by force from a surgical tool at the surface of the liver. Accurate displacement depends on accurate biomechanics. The liver is a mechanically complex organ. It is heterogeneous, being comprised of a spongey parenchyma permeated by a branching network of blood and bile vessels. Each tissue type has its own independent stress strain relationship, for example cancer is 5-28 times stiffer than the background tissue [64]. The biomechanical model must be simple enough for video rate simulation, yet accurate enough for clinically useful landmark placement. A popular modeling strategy is to start with “Atlas” data from experimental observations of the stress strain curve to model biomechanics, [65, 66, 67] and then simplify down to a linear model which can be simulated quickly. This strategy is well suited to video rate simulation speed and it is assumed that modeling prior direct observation of the biomechanics can mimic the liver closely. However, this assumption can break down. Liver stiffness changes significantly with disease and varies depending on the tissue type [68]. Some research exists for applying atlas biomechanics to the patient specific anatomy of landmarks such as vessels and tumors [69, 70]. However, the fundamental issue of patient specific stiffness has remained an open problem in the CAIG community. It is not possible to create patient specific stiffness maps for this preoperative step using the stress relaxation methods used to build an atlas, since the experiment would have to be done intraoperatively, paradoxically when CAIG is needed.

Build Simulator Model

The Simulator model combines liver biomechanics and geometry into a physical model that deforms in response to forces from surgical tools. In CAIG the model is used to describe how the liver geometry model of visual landmarks have been positioned by tool forces applied to the liver. Complex liver models use specialized elements to model specific tissue types such as shell elements for the Glisson's capsule and beam elements for large vessels [71]. Mesh free models are also possible. These particle based approaches have benefits when modeling the heterogeneous materials of the liver since the meshless approximation allows a smooth transition of material properties at interfaces between the parenchyma, vessels and capsule without exhibiting sharp discontinuities at the interface boundaries [72, 73]. It is key that liver models are representative of liver tissue. While much work in CAIG modeling has focused on efficient phenomenological consitutive models and heterogeneous discrete element compositions [74], the state of the art does not consider patient specific liver stiffness. Surgeons operate on diseased livers and liver stiffness is known to vary with disease [75]. Because the liver is large and heterogeneous, CAIG that does not use patient specific stiffness data is likely to be inaccurate. To the best of our knowledge, *no CAIG method can accurately model patient specific liver stiffness.*

Camera Calibration

This step is a prerequisite for mapping between intraoperative laparoscope video and landmark geometry models. Camera calibration finds a transform that is used to calculate the distance and size of objects seen by the laparoscope in real world units [76]. Camera calibration is performed routinely since stereo laparoscopes provide a surgeon with depth perception in robotic minimally invasive surgery. Without calibration, misalignment of the cameras and lens distortion confuses human depth perception [77]. The calibration parameters describing lens distortion are referred to as the Intrinsic matrix. The parameters describing relative pose of the stereo camera pair are referred to as the Extrinsic matrix. Both can be combined into a camera matrix. Camera calibration is well studied and methods have been proposed to automate the computation of a camera matrix using a set of images [78, 79]. It is more important to calculate the focal length accurately in CAIG applications than in traditional surgery. This is because online steps must calculate movement in the scene in real world units and a poor camera matrix reduces the accuracy that can be achieved [80]. Since this step is offline, traditional pattern based calibration methods can be performed routinely. Sub-millimeter calibration precision has been demonstrated on standard flat facing and oblique-viewing laparoscopes [81, 82] at 30Hz, which is an acceptable clinical standard.

Initial Registration

An initial registration defines the first transform between the surgical scene and preoperative scene model. In other words, it definines the initial position of image guidance landmarks at the start of surgery. This step bootstraps subsequent registrations and is sometimes refered to as “hand-eye” calibration in the robotics literature [83]. In general, a rigid object of known dimensions such as a calibration cube is placed in the camera field of view. A virtual camera is defined in the simulator coordinate system and this virtual camera is pointed at a virtual version of the calibration object. An initial registration transformation matrix can then be computed via linear decomposition using the appearance of the object in both coordinate systems. In the CAIG context, the calibration object analogue is the liver. Despite the readily deformable nature of the liver, at the start of a procedure the patient is often in the same position as when preoperative scans were taken. Rigid and semi-regid registration methods have therefore shown promising results [84, 85]. Alternatively, the CAIG pipeline simulator can be used to compensate for non-rigid shape changes during initial registration [**?**, 86, 87]. The surgical environment poses other challenges besides initial discrepancies in liver shape. For example, much of the liver is outside the laparoscope field of view in minimally invasive procedures. Although this makes colocating salient features in the laparoscope and simulation model difficult, semi-automatic initial registration has been demonstrated with either 30% of the liver visible, or an intelligent choise of laparoscope viewing angle [88, 89]. Good quality automated registration has been demonstrated when landmarks can be reliably found in the preoperative and intraoperative images [90]. However, the initial registration step typically must be performed manually or use semi-automatic methods to reach acceptable clinical standards [91, 92, 93].

Camera To Robot Registration

In this step, the real world surgical tool positions are mapped to virtual tool positions. The virtual tools must be positioned in the same coordinate system as the image guidance landmark model. Then, when the real tools touch the real liver, virtual tools will touch the virtual liver and these interactions can be modeled by the CAIG pipeline. This step is sometimes a part of hand-eye calibration since the goal is to compute the initial transform of tool models relative to the initial registration. Camera to Robot Registration is usually less challenging than the initial registration because most tools are rigid, which makes calculating the transform simpler. Conventional hand-eye calibration methods can therefore be applied more easily. These include edge feature methods [94], local gradient descriptors [95], or color features [96]. Tools are not patient specific, so accurate CAD models can be created from specifications. when tools are seen in the surgical scene, the accurate tool model created offline can be used for hand-eye calibration directly. Examples of this approach include 3D and 2D object based hand-eye methods that use tool models as the known calibration object [97, 98, 97, 99].

3D Reconstruction

This step converts flat 2D laparoscope images into a 3D representation. This is the first *online* step in fusing the intraoperative video to 3D landmarks in the liver geometry computed *offline*. Typically, 3D reconstruction is done on stereo laparoscopes, but monocular laparoscopes are also common. Liver surgery is a challenging environment for traditional optical reconstruction approaches that rely on stable features [100]. These methods are not robust to confounding factors like smoke and specular highlights, which can easily overwhelm feature detectors with noise. Another issue is that 3D reconstruction is computationally expensive and the step must run at video rate to be of clinical use. Specialized methods have been developed over the past two decades to solve these challenges. The earliest works used FPGA hardware acceleration of computational bottlenecks with clinically acceptable frame rates of 10Hz being reported by Su et al. in 2009 [101]. Systems achieve speeds of 200Hz on laparoscope resolutions of 960×540 or higher with contemporary GPU acceleration and algorithms [102]. A summary of video rate 3D reconstruction for laparoscopic surgery and a survey of their performance is given by maier et al. [103, 104]. Passive stereo or monocular laparoscopes have limited potential for accurate 3D reconstruction and would preferably be replaced. Active laparoscopes are a desirable alternative. This type of laparoscope encodes a ground truth signal into the surgical scene that allows more precise 3D reconstruction as the signal is read back. Structured light laparoscopes are a promising active laparoscope type that project a well characterized light pattern into the surgical scene. Traditional hand-eye techniques are then used on the pattern appearance to reliably construct a 3D depth map. Sui et al. provide a recent survey of the state of the art while reporting a 40,000 point structured light laparascope system with 0.2 millimeter error that runs at 12Hz [105]. This performance is within the sub-millimeter accuracy possible with surgical robots, and runs at clinically acceptable video rate for CAIG.

ROI Search

The second online step, Region Of Interest (ROI) search, keeps the laparoscope camera position synchronized with the virtual camera position. It is used to keep track of which portion of the liver geometry is visible when the laparoscope is moved.

When the laparoscope alone is used to perform the ROI search, this step is a form of the "Simultanious Localization and Mapping" (SLAM) problem. SLAM techniques seek to track the position of a vehicle (in this case the camera) as it navigates an environment (in this case the surgical scene). Whist the camera position is being tracked, SLAM builds a geometric map of the scene [106, 107]. SLAM methods are also online which is essential for the ROI search step. However, traditional SLAM is not well suited to CAIG. Most SLAM assumes a static environment, but the surgical scene is not static due to the patients respiration and a because the liver moves during surgery. Specialized surgical SLAM variants have been proposed to tackle this problem, but key challenges remain. Lin et al. provide a recent survey of surgical SLAM methods that are robust to motion in the surgical scene that summarizes the open problems [108]. It remains challenging to define a deformation model that can be used to filter non-rigid motion in the surgical scene and no optical method has been proposed that can meet clinically acceptable standards.

When the laparoscope is instrumented for motion tracking the problem is less challenging. Fiducial markers are a popular choice and can be applied to the laparoscope and tracked optically or electronically. Xiao et al. recently compared state of the art hardware in optical and electromagnetic tracking [109]. When marker detectors for either method can be placed within 250 millimeters of the proximal end of a laparoscope, high accuracy is possible. Optical trackers achieved 0.210 millimeter accuracy at 335Hz and electromagnetic trackers achieved 0.367 millimeter accuracy at 40Hz. Provided that the laparoscope can be instrumented, accuracy and video rate requirements for clinically acceptable CAIG can be met by using tracking markers.

3D Registration

Non-rigid 3D registration is the first step in tracking displacement of surgical landmarks. It is an online step used to keep track of liver surface motion, which is necessary to track surgical landmark movements. This step must determine the deformation model, usually assumed a brownian warp [110], that maps liver surface in one laparoscope frame to the next. This is a computationally intensive optimization problem that can be made more challenging by inaccuracy in previous online steps (3D reconstruction, ROI search). The main computational bottleneck is estimating the warp field. Unlike rigid registration that computes a single transform, non-rigid registration must compute a transform per feature point. A popular approach is to use parametric flow fields to define a global relationship between features and minimize a global set of residuals [111]. In this approach, the quality of results using energy minimization are highly dependent on: The quality of optical features, the quality of the cost function, and the number of optimization iterations. Optimization based methods have significant overlap with the laparoscopic SLAM approaches mentioned in the *ROI Search* discussion [108]. They suffer from the same challenges of accuracy and drift and are difficult to implement at video rate. Recently, song et al. proposed a GPU accelerated 3D registration method applicable to liver CAIG [112]. Although good frame rate was reported, the method was not robust to large deformations or rapid camera motion and required expert manual selection of key points. A second approach is data driven, where empirical observation of examples is used to estimate the deformation model. Data driven methods can be trained and optimized for video frame rates using machine learning techniques without explicitly defining an optimization function. Pfeiffer et al. demonstrate a GPU accelerated method that achieves close to 5 millimeter accuracy at 50Hz [113]. However the method was trained and validated on synthetic data from physical simulations, making the clinical merits of the method unclear. Despite recent progress, no online non-rigid registration method has been demonstrated as meeting clinically acceptable standards.

Inverse Kinematics

Surgical tools apply force on the liver, displacing internal landmarks. Force is typically computed as input to the physical simulator using active electronics. One approach uses strain gauges mounted to the tool surface [114, 115] however cutting or cauterizing surfaces cannot be easily equipped with sensors. A more flexible approach is to use the electrical resistance of the robot arms and formulate an inverse kinematics problem to determine the position and load condition of the tool tip [116]. Because of factors such as mechanical tollerance and multiple points of articulation, tool force estimates do not approach the precision of pressure sensors. Optical methods are a third option for estimating inverse kinematics [117], however tool priors such as dimension and mass must be known for quality results. Because of this, the same techniques applied to compute camera to robot registration can be applied to this step, including pose tracking methods [98].

Physical Simulation

The online physical simulation step is used to position surgical landmarks in response to the force acting on the liver estimated in the previous inverse kinematics step. Interactive physical simulation in CAIG must strike a balance between mechanical realism and video rate performance. It is challenging to make the right tradeoffs. A recent survey of surgial simulation methods by Zhang et al. describes state of the art methods [118]. Finite Element Method (FEM) simulation is popular in the literature. It is accurate, but also computationally expensive [119]. Several FEM algorithms have been developed for image guidance applications that can model soft tissue deformation with video rate speeds on CPU [120]. These algorithms address the problems caused by large deformations of soft tissue such as the liver, which can break the displacement assumptions of standard elements. Specifically, computationally efficient solutions are given for problems such as element locking and volume hourglassing. More recently, hardware accelerated FEM for surgical simulation applications has been demonstrated on GPU [121]. This method implements an anisotropic visco-hyperelastic constitutive material model, which has applications in liver modeling. Other simulation methods besides FEM have been applied to the problem, although these are less common. Meshless methods have some advantages in surgical situations, such as being more suitable for representing the large deformations of the liver, and modeling cutting without the need for remeshing. However, the particle based nature of meshless models makes handling object boundaries challenging, often relying on an interpolation function that reduces accuracy. A prominent example of meshless simulation is De et al. who describe a liver surgery application of their finite spheres method [122, 123]. More recently, machine learning [124] and data-driven simulations [125] have been proposed. These methods are usually not iterative and are typically accelerated with GPU for high speed. Machine learning and data-driven methods have potential for good accuracy, but require high quality databases of empirical experiments. In general, although much progress has been made in video rate simulators, the accuracy of the physical simulation step is highly dependent on good quality liver models created earlier in the *Build Simulator Model* step.

Augmented Display

This final online step presents image guidance landmarks to the surgeon. Figure 1.1 illustates an “Augmented Reality” video overlay method where landmarks are added to the laparoscope video frame [126]. Once surgical landmarks positions have been computed, the CAIG overlay can be rendered. Compared with prior online steps, augmenting the display is not computationally intensive. First, a mapping of the geometric model of guidance landmarks from the simulatior coordinate space to laparoscope camera space is applied. This map is the inverse transform that was computed during the *ROI Search* step. Next, surgical tools may be segmented from the laparoscope video overlay so that surgical tools appear to be rendered in front of the image guidance landmarks. The segmentation is done using tool masks computed during the *Camera To Robot Registration* step. Surgical landmarks can then be rendered as superimposed anatomical structures on the laparoscope surgical video feed [127]. Although many visualization methods have been propsed [128], the clinical merrits of most are unclear. Kersten-Oertel et al. provide a good taxonomy of surgical visualization approaches that higlights a trend of lapse validation in the literature [129]. Ideally, augmented display research focuses on the ergonomic challenge of presenting the landmark overlay in a way that does not distract the surgeon and lead to poor surgical outcomes [130]. However, it is uncommon for work to focus on this topic alone. Landmark positioning does not meet clinically acceptable standards due to the open research problems of the earlier *3D Registration* and *Physical Simulation* steps [131]. Thus, progress in augmented display research is highly dependent on progress in these earlier online steps.

1.6 Research Problems Of Focus

As outlined in the previous section, the online steps of 3D registration and surgical landmark placement have disproportionately hampered progress towards clinically useful CAIG. We therefore focus on efficient registration and landmark tracking in this manuscript. Improving efficiency means improving accuracy without cost to speed or speed without cost to accuracy in the CAIG pipeline step. In particular we make contributions to efficient online algorithms for tracking the surface of the liver via non-rigid surface registration, and to image guidance landmark positioning in liver surgery via physical simulation. Below we describe the specific challenges that we overcome in this manuscript to make our contributions.

Non-rigid 3D Registration for Surface Motion Tracking

Tracking surface topology changes in laparoscopic surgeries is a hard computer vision problem that must be solved online in real time. There are two major challenges. The first is identifying many salient features and tracking them correctly throughout the video. Feature trackers often perform poorly with surgical video as the surgical environment is fraught with confounding visual artifacts such as glare, specular highlights, partial occlusions caused by bodily fluids, tools, or gross anatomy, poor lighting, and blur. The combination of these challenges is highly nuanced depending on the type of surgery as well as the type of surgical tools.

The second major challenge is registration of salient features once they have been reliably captured. The liver is extremely malleable and its surface smoothly deforms when force is applied to it. Registration of the features therefore requires finding a smooth transform across all features that properly describes the continuous deformation of the surface for all features between consecutive video frames. Ideally, the transform should have as many basis functions as the number of tracked features, be continuous and be everywhere differentiable to allow for interpolation of the transform for non-feature points on the surface. Accurately solving this optimization problem is an open research area since it is not trivial to define the appropriate optimization framework or the associated cost functions. CAIG also has the additional real time performance constraint which adds another dimension of complexity to the challenge since any registration framework must run fast enough to keep pace with the video frame rate.

Presently it is not tractable to implement global registration for a large number of features and the literature focuses on various tradeoffs to meet real time performance goals [132, 133, 134]. It is not trivial to evaluate which trade-offs are best suited to clinical applications. Registration trade offs are strongly affected by the quality of features and algorithms are often not designed for surgical applications.

We therefore propose that inaccuracies of CAIG are partly caused by a lack of surgery centric video registration evaluation metrics. However as mentioned, accurate registration alone is not enough. The registration step of fig. 1.1 must also be performed at video frame rate so that the following step can use this information to position IG landmarks for the surgeon in real time.

Physical Simulation for Surgical Landmark Positioning

Despite the existant understanding that physical simulation can model biomechanical properties to provide accurate CAIG in liver procedures [135, 136], the necessary tissue mechanics are complex to analyze in vivo and our present view of accurate CAIG liver resection is quite narrow.

Existing CAIG works assume that the liver biomechanics are defined primarily by the geometry of the organ, and in particular the vasculature [70, 69]. Recent studies have suggested biomechanical properties of the liver are associated with disease stage and primarily heterogeneous in the parenchyma [137, 138] as opposed to the vasculature structure, which suggests that CAIG requires patient specific mechanical models. Furthermore, existing CAIG approaches require manual tuning of stiffness to reach convergence with the observed liver during surgery, which demonstrates that conventional stiffness assumptions and biomechanical models are insufficient for accurate CAIG in vivo [70].

Thus we propose that current inaccuracies of CAIG is caused by patient specific heterogeneous variation of mechanical properties in the liver. Despite the alarming increases in liver cancer incidence and morbidity, preclinical and clinical research has provided little progress developing effective CAIG with acceptable accuracy margins. Consequently, all operability decisions are made assuming an order of magnitude less accuracy in robotic tools than they are capable of.

However, the accuracy of CAIG cannot be improved without consideration of other important factors. The implementation of liver modeling algorithms must take into account the complexity of the offline modeling steps and the power consumption of the online simulation steps. Without efficient implementation, advanced liver models are not practical for video rate CAIG.

1.7 Summary Of Contributions

We contribute several data driven efficiency improvements to the canonical CAIG pipeline. Specifically our methods improve efficiency of *non-rigid 3D registration* for liver surface motion tracking and *physical simulation* for image guidance landmark placement. Highlights include a registration algorithm and hardware accelerator that is more power efficient and robust to large changes of liver shape than the related work. We also describe a physical simulation method that is the first to model heterogeneous patient specific liver stiffness. Using this data to place surgical landmarks is more accurate than the related work and has no computational overhead in the online step. Our methods demonstrate new techniques applicable to open research challenges and are valuable contributions towards practical CAIG in liver surgery. We detail our contributions as follows:

In Chapter 2, we demonstrate how SIRGn makes comparing the accuracy of registration algorithms on laparoscopic video simple compared with the traditional "Target Registration Error" and "Root Mean Squared Error" metrics.

Chapter 3, we describe the "TRWL-S" non-rigid registration hardware accelerator and demonstrate how it improves the efficiency of an accurate registration algorithm chosen using a SIRGn analysis. we demonstrate how high level Hardware Descriptive Language (HDL) can be used to rapidly implement an efficient hardware accelerated version of the most accurate registration algorithm chosen in Chapter 2 for the registration step. We apply a bevy of hardware software co-design methods to develop a hardware friendly version of the registration algorithm and find a large speed up with a minimal accuracy trade off.

In Chapter 4 we consider the simulation step. We Introduce the idea of improving CAIG simulation accuracy by using Magnetic Resonance Elasography (MRE) to create patient specific data driven physical simulations for positioning IG landmarks. We motivate our proposal with a patient cohort study, explore the accuracy implications of MRE data, and provide an in-silico validation of our approach.

Next, we consider the important factors of modeling complexity and power efficiency for patient specific landmark placement. Our MRE results in chapter 4 are promising, however, MRE is an exotic imaging modality that is not accessible to many surgeons.

In chapter 5 we consider less well equipped hospitals. We propose a simple alternative data driven simulation approach to patient specific landmark placement that does not rely on MRE. Our method uses a series of empirical measurements taken using a cheap strain gauge and a novel learning based physical simulation for patient specific CAIG landmark placement. Surgical robots are also somewhat uncommon. They are expensive platforms that may not be available to resource limited surgeons.

In Chapter 6 we consider the situation where CAIG must be implemented on commodity power constrained hardware such as laptops or mobile phones worn by the surgeon during liver surgery. We propose a framework for clinicians to tune efficiency of CAIG using case specific parameters on low power mobile devices. We begin by showing the bulky size of laptops can create a practicality challenge to clinicians. Namely, they easily get in the surgeons way and restrict movement around the patient in the operating room. Next, we use a representative set of contemporary mobile hardware acceleration devices to define a design space ranging from cumbersome laptop back packs to head mounted platforms. Finally we define a platform design framework for clinicians to find an optimal CAIG platform. We demonstrate how the case specifics of a liver procedure can be used to find the least obtrusive CAIG platform for that procedure.

Chapter 2

SIRGn

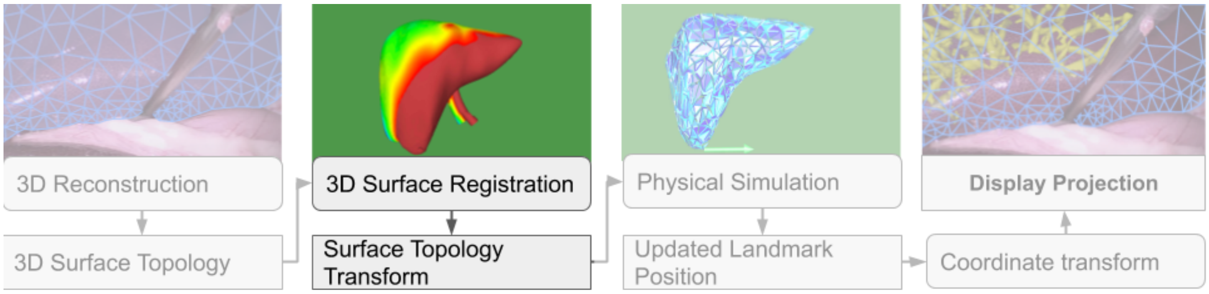


Figure 2.1: SIRGn is concerned with improving the accuracy of the online 3D Surface Registration step in a CAIG pipeline.

Recall the four major online steps required for accurate CAIG outlined in chapter 1. In this chapter we describe SIRGn, a method for improving the accuracy of the 3D surface registration step. Figure 2.1 highlights the registration step in the CAIG pipeline to illustrate where our contribution is useful within a larger CAIG system.

2.1 Introduction

Augmented Reality (AR) has significant potential for enhancing surgical procedures. For example, the ability to superimpose landmarks like major internal vessels onto the surface of an organ provides immense value for surgical image guidance. There are many therapeutic uses for AR guidance in laparoscopic surgery [136, 139, 140, 141, 142] and other closely related intervention types such as telerobotic surgery where AR enhances the pre-existing feedback from robotic tools [143, 144].

A key challenge for AR surgery is registration, which provides the relationship between a preoperative model (e.g., CT or MRI scan) and the surgical video feed. Fig. 2.2 illustrates the registration process. The three images/frames on the left are collected in real-time during the procedure from a camera. Each frame must be registered onto the preoperative 3D model on the right. After that, features from the preoperative 3D model are overlaid onto the surgeons view.

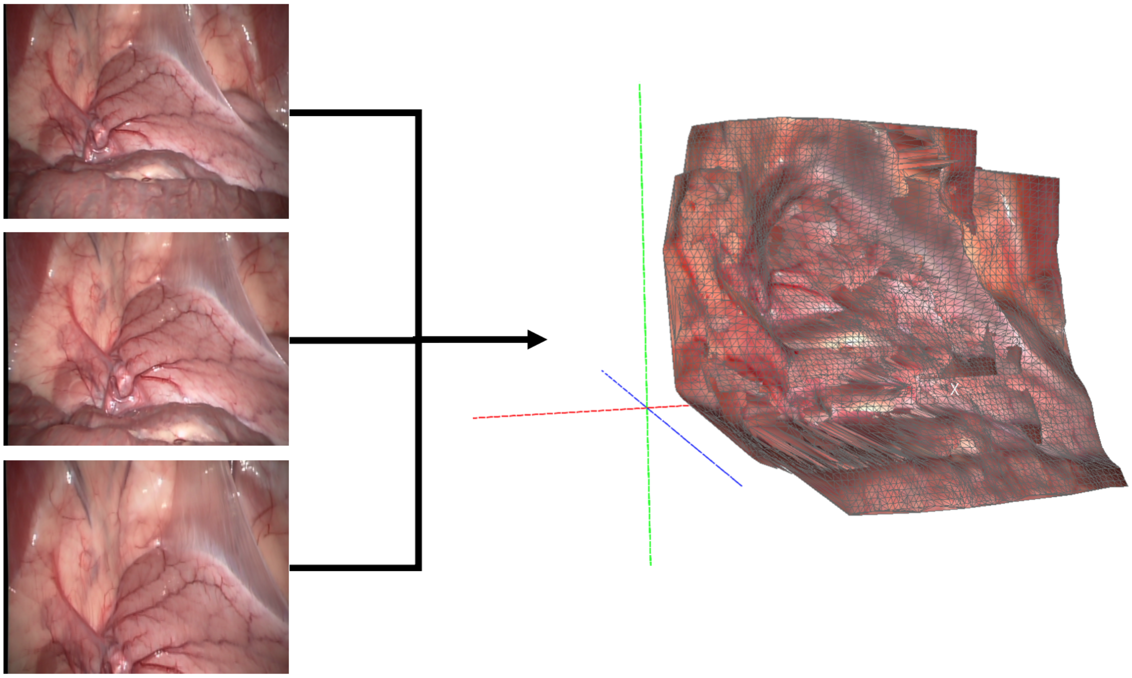


Figure 2.2: Registration maps video frames (left) onto a 3D model (right). The video frames should be mapped in real-time onto the 3D model allowing any features from the model to be projected into the surgeons view.

Surgical augmented reality is particularly challenging since it must be done in real time with high accuracy. A surgical scene can be extremely complex: tissues deform, lighting varies dramatically, and features quickly come in and out of the field of view. Every laparoscope, every lighting solution, and every type of surgery have unique idiosyncrasies. Consequently, custom registration algorithms are needed for the unique challenges arising from a given combination of factors in a particular surgery. It can be difficult to tune the performance of a registration algorithm to the surgeon's needs since benchmarks that account for specific idiosyncrasies of that surgery do not exist. Because of these challenges, surgery-specific heuristics are needed to perform high quality dense surface registration in real-time. However, it is impossible to measure the effectiveness of these heuristics without a baseline quality measure. A useful quality measure must be *dense, granular,* and *convenient*.

Many quality metrics have been applied to clinical registration. The canonical approach is Target Registration Error (TRE) which involves computing $L2$ or some other distance between corresponding feature points after registration [145, 146, 147, 148, 149]. Feature separation distances can be highly representative of registration error if the features are appropriately chosen, but such metrics have no concept of a global coordinate system to track drift in a stream of registrations. Besides distance measures, global similarity metrics based on intensity have broad clinical applications [150, 151, 152, 153]. These qualitative measures combine intensities of the registered images into one image. A clinician can then judge the quality of registration through experienced observation of the resulting overlay of images. Because of this reliance on clinical experience, intensity metrics are not used in applications where registration accuracy is critical as they cannot represent the warp function in a physically meaningful way.

Landmark registration metrics judge registration quality by an objective function weighted by pre-defined feature points. These consider rigid registration [154] and/or fiducial markers [155] rather than tissue only landmarks and can be used to focus the metric on registration regions of clinical interest. Hoffmann et al. [156] develop such a quality metric. Unfortunately, they do not attempt large surface area coverage, they rely on a small number of landmarks, and their technique is applied to CT scans rather than video. Recently Thompson et al. [157] proposed an AR specific registration evaluation focused on accurately positioning sub-surface landmarks. Unfortunately shifting focus from surface registration to sub-surface features requires that assumptions are made on the biomechanics of the organ when estimating landmark positions. Because such assumptions often break down [75], quantitative surface registration quality metrics are more robust.

Surgical Image Registration Generator (SIRGn) is a novel dense, granular baseline quality metric for video registration algorithms. Fig. 2.3 shows how SIRGn works. It evaluates the registration of a set of images/frame (Fig 2.3a) by creating a heat map that shows regions of poor registration quality (Fig. 2.3b) that can be subsequently overlaid onto the scene (Fig. 2.3c) for visual inspection of the quality of result of the registration algorithm.

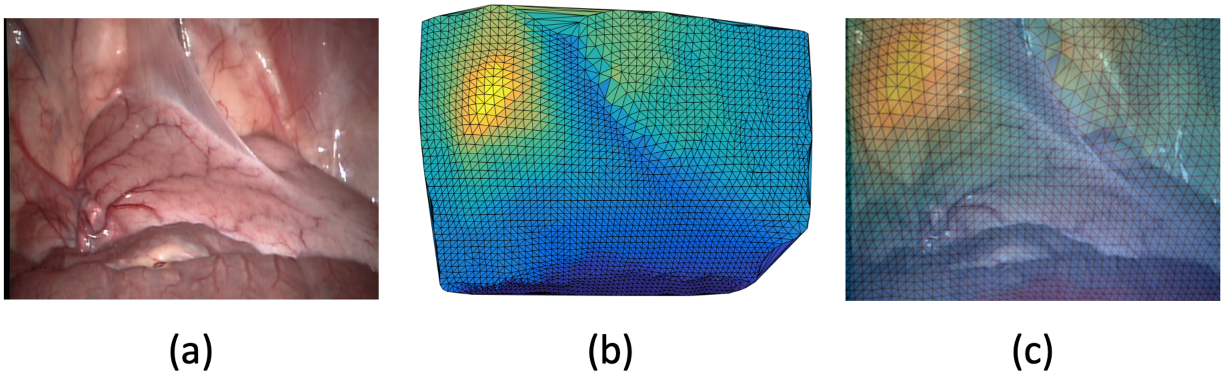


Figure 2.3: Illustration of the SIRGn concept. Assume a 3D registration algorithm is run on the video frame shown in Fig. 2.3a. SIRGn evaluates the quality of the registration by creating a covariance mesh (Fig. 2.3b) and overlaying that as a "heat map" (Fig. 2.3c) where high heat indicates poor registration quality, e.g., as seen in the upper left corner.

The novelty of SIRGn lies in the fact that it combines the physically meaningful metric of TRE with the clinically useful dense image coverage of an intensity type measure. Additionally, unlike the aforementioned approaches, SIRGn is well suited for image guidance benchmarking since real time registration algorithms are evaluated using a global coordinate system, which allows it to compare across multiple frames. Our method can be applied in situations where no suitable database of videos exists or where synthetic simulation is intractable. Also, it can be applied to many types of surgery and it does not rely on physical simulation or biomechanical tissue information.

In this manuscript we describe the SIRGn method and an experiment that uses SIRGn to compare performance of two non-rigid 3D registration algorithms on laparoscopic video from a database [158]. Our experiment models the exploratory phase of AR image guidance where several approaches are tried to accurately track the clinically significant portion of a surgical scene over time. Specifically we show that SIRGn is advantageous in evaluating the accuracy of different registration methods by comparing a SIRGn work flow with the conventional TRE and RMSE error metrics. Registration performance is evaluated using landmark correspondences [159]. Although it is desirable in clinical image registration to use expert landmark correspondences, our goal is convenience[1] and so we provide tools for automated global reference feature extraction from 2D and 3D laparoscopic video.

The remainder of this paper is organized as follows. Section 2 gives an overview of the baseline generation method and evaluation of 3D registration quality using SIRGn as a baseline. Section 3 gives an appraisal of SIRGn as a simple and flexible baseline generation method. We conclude in section 4.

2.2 Method

SIRGn provides a way to evaluate a given registration algorithm on a surgical 3D video data set. For this, we must create a metric that describes the quality of the registration. An ideal metric would work for any registration algorithm and data set. It would allow us to compare the quality of different registration algorithms on that data set. And it would provide a location

---

[1] Example code and data available at: https://github.com/KastnerRG/SIRGn.git

specific quality measure. We describe our metric in the context of stereo laparoscopic cameras since we anticipate SIRGn's greatest applicability to be in image guided surgery. However our methodology can be applied to conventional 2D surgical video data.

The insight motivating SIRGn is that surfaces in surgical scenes are smooth with very few exceptions. Therefore given a smooth surface, a point that is heavily warped by registration with respect to surrounding "known-good" points is less likely to be accurate. In order to formalize this intuition, we rely on two sets of points. The first, G, is a set of expert landmarks that we assume as correct a priori; these could be human labeled by a surgeon or provided by a reliable feature detection algorithm. The second, C, contains points that require heavy warping to align well on a global model M. Points in C are chosen by searching points enclosed within the triangulation of G. Specifically, a set of candidate points to be registered is masked using a triangle from G to project onto M space. After registration the point that required the largest registration warping transform to M space is added to C. Any point that requires substantial warping is likely due to an incorrect registration, and similarly any points that require a small amount of warping are likely correctly registered. Thus, the key is to determine how to evaluate these warp functions for any registration algorithm and to derive a method to efficiently locate points C that must be heavily warped. We describe this in more detail in the following.

The four major components to SIRGn are shown graphically in Fig. 2.4. Fig. 2.4a computes a non-rigid warping function $\mathscr{S}$ that projects G from the video frames V to the 3D model M. That is, the warp function $\mathscr{S}$ describes the relationship between the landmarks G in the stereo video frames V and those same landmarks on the 3D model M. One can view M as containing the average location of all the landmarks within G over time. Fig. 2.4b computes a Delaunay triangulation between each landmark g_i on the 3D model M. Fig. 2.4c finds the most heavily warped points for each surface triangle. Fig. 2.4d assigns a value to each triangle computed from these sets to faces of a Delaunay triangulation to form a "correlation mesh". This is the output that indicates how well the worst case points correlate with the landmarks. If

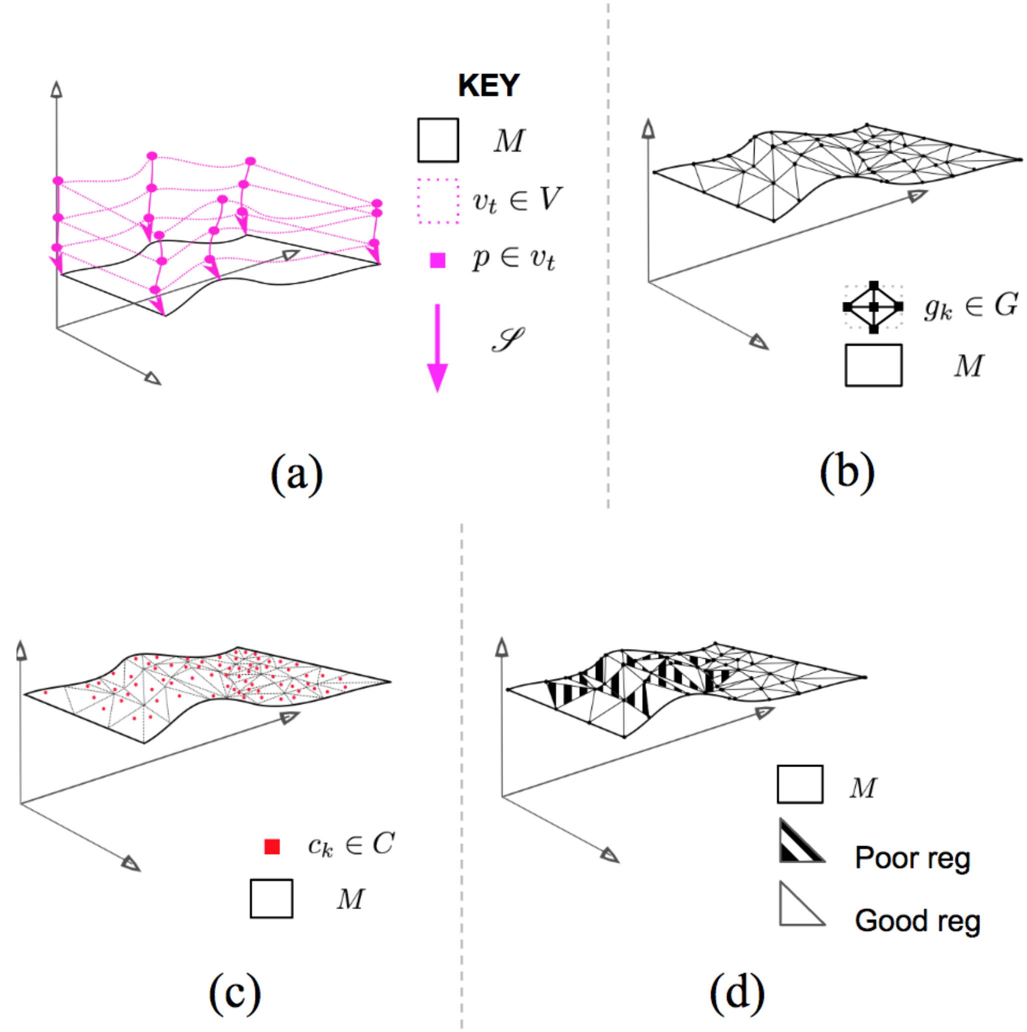


Figure 2.4: Steps to compute the SIRGn quality metric are, in order: (a) Construct a 3D model M and warp $\mathscr{S}$ using points $g_k \in G$ and all 3D scans $v_t \in V$. (b) Project expert-labeled landmarks $g_i \in G$ onto M. (c) Find heavily-warped registered points $c_i \in C$. (d) Label a mesh with vertices G showing registration quality. The labels/coloring are a function of G, C and ε tolerance and can be visualized, e.g., as done in figs. 2.6d to 2.6f.

there is high correlation, then the worst case points correspond well to the assumed high quality landmarks, and thus are more likely to indicate that the registration was accurate. This correlation mesh is very useful for comparing the performance of different registrations with one another since it highlights regions of relatively poor performance.

Once a candidate algorithm has been selected using the correlation mesh, SIRGn can be re-run and set to report mean registration error of the G triangulation for a candidate registration in another mesh. We set an acceptable error threshold based on the registration accuracy desired. This is typically defined by the needs of a clinical AR application, i.e., 5mm is typical abdominal surgeries [44]. Unacceptable regions of the candidate registration are highlighted in a second mesh. This work flow is demonstrated in Section 2.3.

2.2.1 SIRGn Registration Quality Measure

In order to use SIRGn, an evaluator must first run a global registration algorithm, which we call ALG. ALG is chosen to be highly accurate, i.e., real time constraints do not apply in order to maximize the registration quality. Because AR registration is performed in $\mathscr{R}3$, SIRGn should be run after a stereo reconstruction algorithm on stereoscopic laparoscopes although it can be run directly on 3D scopes. ALG returns a dense global registration model M by performing a global registration optimization that minimizes the transform distance of the set of G points for all 3D video frames $v_t \in V$, such that

$$M = \text{ALG}(V, G) \tag{2.1}$$

Global Warping Function: Once we have computed M, we obtain a dense benchmark registration over M. To do this, we compute a function $\mathscr{S}$ mapping from scans V to the space of M to allow every point in 3D depth scans $v_t \in V$ to be mapped to the dense surface. We choose the method used by Global Non-Rigid Alignment (GNRA) [160] using thin plate spline (TPS) interpolation that is regularized to produce a smooth function.We prefer this method as a warp baseline both for its stability and the guarantee that $\mathscr{S}$ is a smooth bijective transform. The resulting $\mathscr{S}$ is the global warping function, such that $\mathscr{S}(p;t) = \bar{p} \in M$ for any point p in a 3D scan v_t. $\mathscr{S}$ is therefore useful when building the C set as detailed later.

Landmark distance vectors: SIRGn builds a set $\mathcal{D}_G$ of landmark distance vectors consisting of one vector ℓ_k for each expert landmark in the set of all landmarks G. Note that landmark points g_k in some 3D scan v_t correspond to each other throughout scans $v_t \in V$, and to one $\bar{g}_k$ in the space of the global model M. We compute index t of these vectors ℓ_k as the distance between $g_k \in v_t$ and $\bar{g}_k \in M$. That is,

$$\ell_k[t] = \|g_k - \bar{g}_k\|_2, \text{ s.t. } g_k \in v_t \tag{2.2}$$

for each global landmark $\bar{g}_k \in M$. Then $\mathcal{D}_G = \{\ell_k : 1 \leq k \leq |G|\}$ is the set of landmark distance vectors.

Warping distance vectors: SIRGn computes a mesh of M via Delaunay triangulation using G as mesh vertices. Then, for each triangular face Δ_{c_k} of the mesh we define points $c_{k,t}$ in frames v_t as the *most heavily warped in* Δ_{c_k} at time t. Formally,

$$c_{k,t} = \operatorname{argmax}_{p \in \Delta_{c_k}} \mathscr{S}^{-1}(p;t) \tag{2.3}$$

Note that since $\mathscr{S}$ is bijective and everywhere differentiable, $\mathscr{S}^{-1}$ is well-defined for each t, and Eqn. (2.3) can be solved using standard optimization techniques. We choose these points $c_{k,t}$ because a point on the model surrounded by expert landmarks on the surface of an organ that undergoes a different warp than these landmarks is unlikely to track the tissue surface well. We form a set C_t of these “worst points” $c_{k,t}$, one for each triangle in the mesh, by projecting c_k’s enclosing triangle to form a planar region of interest (ROI) and then performing gradient ascent on $\mathscr{S}^{-1}$ within these ROIs. Once C_t is determined, the set $\mathcal{D}_{C_t}$ of warping distances is computed identically to $\mathcal{D}_G$ with $c_{k,t}$ and $\bar{c}_{k,t}$ replacing g_k and $\bar{g}_k$ in Eqn. (2.2).

Correlation mesh: In order to make our metric dense over the registration surface, we define Delaunay mesh face-values as the linear correlation between vectors $\omega_{k,t} \in \mathcal{D}_{C_t}$, and averages of those vectors corresponding to the vertices of Δ_{c_k} from $\mathcal{D}_G$. That is,

$$\text{SIRGn}(\Delta_{c_k}) = \frac{1}{n}\sum_t \text{corr}(\omega_{k,t}, \hat{\ell}_k) \tag{2.4}$$

where $\mathrm{corr}(u,v)$ is the linear correlation (normalized covariance) between vectors u and v, and $\hat{\ell}_k = (\ell_a + \ell_b + \ell_c)/3$ where $\ell_a, \ell_b, \ell_c \in \mathcal{D}_G$ are the three vectors corresponding to those g_k that are the vertices of Δ_{c_k}, and n is the number of frames containing c_k. Good registration in the local region Δ_{c_k} is indicated by a large correlation between the vectors for that particular Δ_{c_k}. In this way we obtain a quality score that is both *dense*, because the whole model surface is covered, and *granular*, so that each region can be specifically highlighted for easy localization of unacceptable registration errors.

Labeling: This correlation mesh can be transformed into a binary accept/reject by taking a tolerance parameter ε to define what level of error is acceptable. We can also display the linear correlation on the mesh faces directly via coloring for final visualization. Labeling the Delaunay triangle set of G is the final step in generating the SIRGn quality metric.

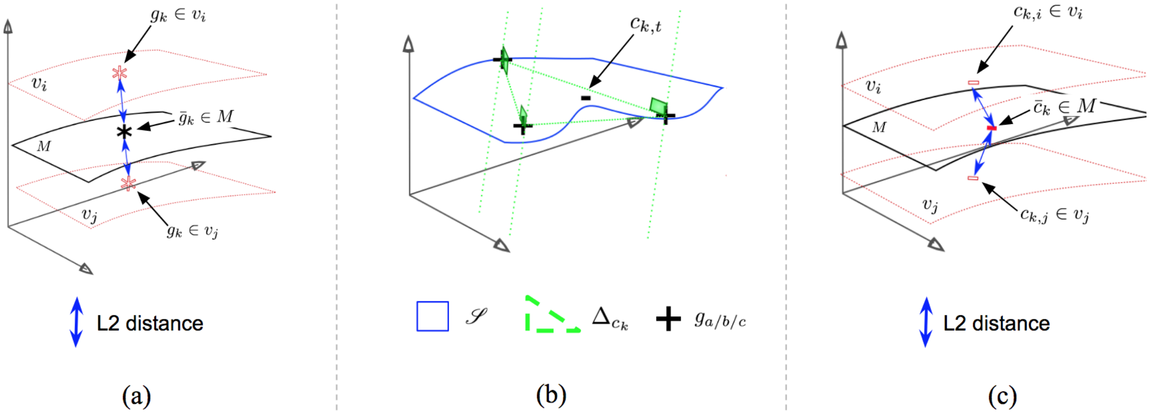


Figure 2.5: (a) Indices i and j of the landmark distance vector ℓ_k corresponding to $\bar{g}_k$ are computed as the Euclidean (L2) distance from the corresponding points $g_k \in v_i$ and $g_k \in v_j$. (b) The most heavily warped point $c_{k,t}$ are computed for a single time point t as the maximum warp $\mathscr{S}$ required to send the point $p \in \Delta_{c_k}$ to the model M. (c) Warping distance vectors ω_k are computed similarly to landmark distance vectors, except $c_{k,t}$ is used instead.

2.3 Experiments and Results

We track respiration motion in a laparoscopic surgery scene accurately in order to best place image guidance landmarks on a live patient. We use SIRGn and the conventional metrics of

RMSE and TRE to compare two 3D surface registration algorithms on laparoscopic video from the Hamlyn data set [158]. Our goal is to use each of the comparison metrics to understand which registration algorithm was best suited to tracking the non-rigid respiration motion in the scene. Our findings detailed in this section showed that SIRGn meshes offered a simple way to compare registration in specific regions of interest in a case where traditional metrics were more difficult to interpret.

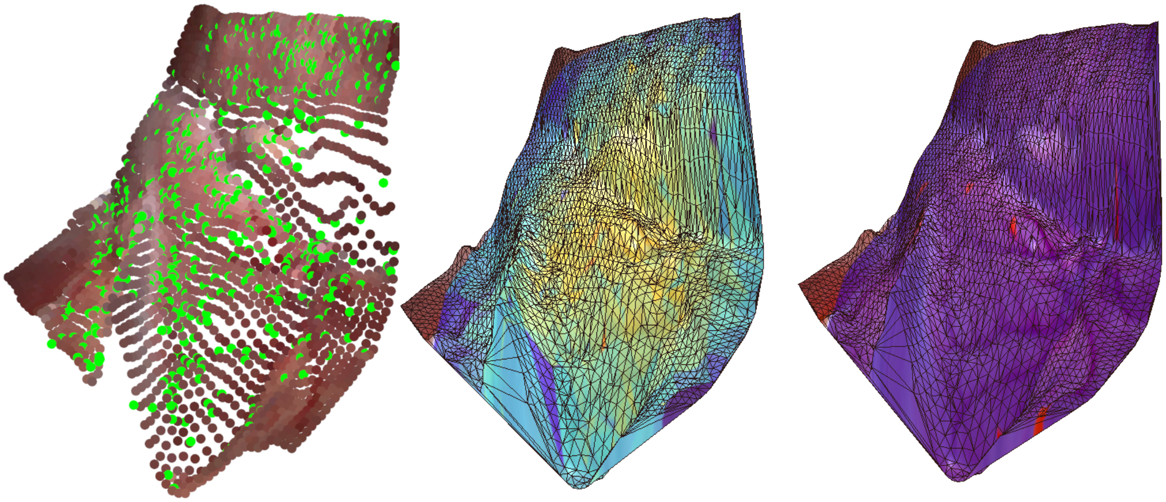
We ran all metrics on two algorithms under test. The first, "$ICP(V)$", is a non-rigid ICP variant based on Amberg's work [161].The second, "$GEO(V)$", uses Chen's geodesic approach [162]. Geodesic registration is suited to non-rigid problems but is typically computed offline due to its complexity. However, a real time implementation was recently described, making it interesting to compare with ICP [163].

SIRGn was run on the algorithms under test using the following steps. First the $ALG(V,G)$ inputs were generated from a laparoscopic video data base [158] using a stereo reconstruction specialized to laparoscopic video (V) [164] and a robust global SIFT tracker (G). Although mixed manual methods for extracting ground truth are more desirable, these must be applied during video acquisition which was not possible in our case [165, 166]. Next, C and M were generated by running $ALG(V,G)$.

Correlation: We applied Eqn. 2.4 to the output from $ICP(V)$ and $GEO(V)$ to generate relative quality meshes as shown in Figs. 2.6b and 2.6c. Yellow is weak correlation with the benchmark and deep blue is strong correlation with the benchmark. The color map of the figures was normalized to the algorithms under test for ease of visual comparison.

Labeling: After using SIRGn to compare $ICP(V)$ and $GEO(V)$ quality meshes, we ran SIRGn with a labeling threshold ε. The ε was set to reject error $\geq$ 10,5 and 3mm as shown in figs. 2.6d and 2.6e and fig. 2.6f respectively. Mesh triangle registration errors of ε and above were greyed out.

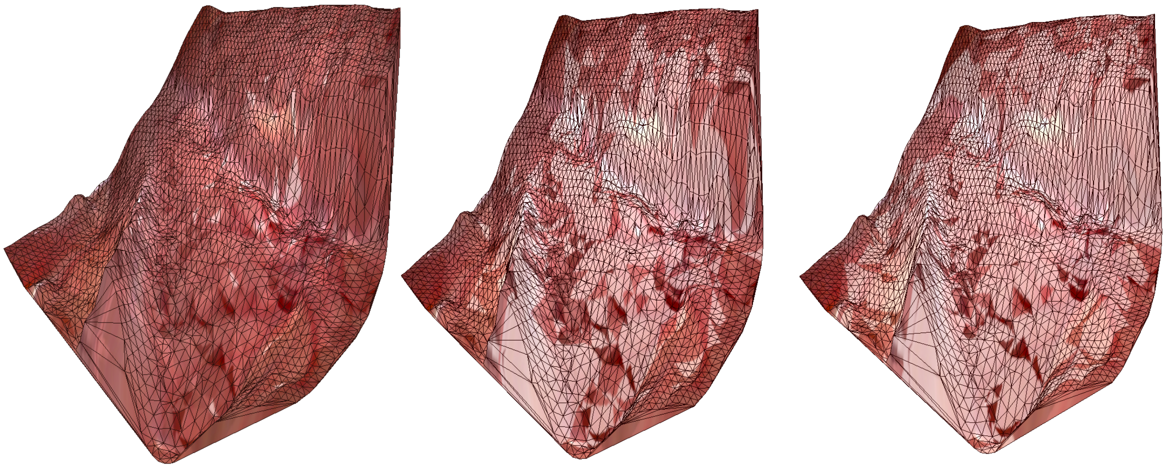
Figure 2.5 summarizes the results of all metrics where $|G| = 1155$ and $|V| = 180$ which



(a) Example 3d V video frame with G points

(b) Normalized ICP registration error mesh

(c) Normalized GEO registration error mesh

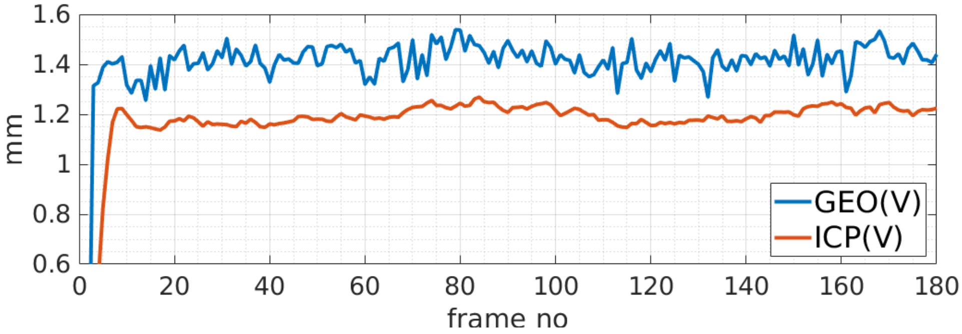


(d) 10mm Threshold GEO registration error mesh

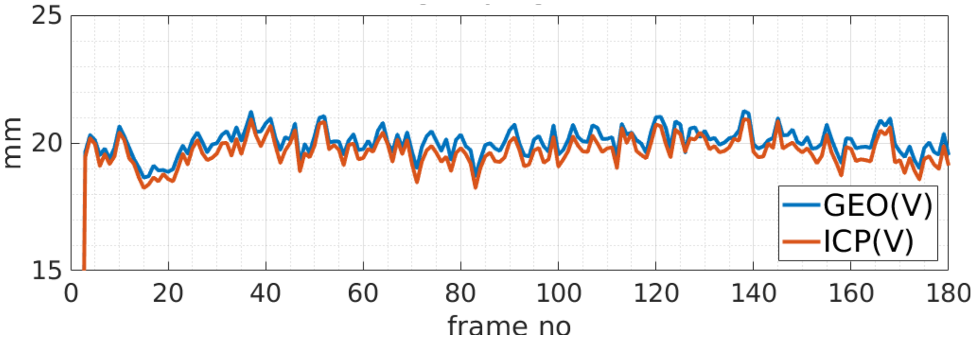
(e) 5mm Threshold GEO registration error mesh

(f) 3mm Threshold GEO registration error mesh

Figure 2.5: Example use of SIRGn to compare two non-rigid 3D registration approaches(figs. 2.6a to 2.6c), alongside traditional error metrics (figs. 2.6a to 2.6c). Figure 2.6a shows a priori known good points (G) rendered on one video frame in green. Figure 2.6b and fig. 2.6c are the quality meshes for the $ICP(V)$ and $GEO(V)$ video registrations respectively. According to SIRGn, $ICP(V)$ showed a far greater variation than $GEO(V)$ from the baseline in the region of video with most movement. Figure 2.6d, fig. 2.6e and fig. 2.6d show the results of SIRGn set to color the $GEO(V)$ error mesh white. based on different triangle registration error thresholds. Important parts of the scene can be inspected at the desired threshold setting to see if registration is acceptable.



(a) Per frame RMSE



(b) Per frame TRE

Algorithm	**RMSE**	**TRE**
GEO(V)	1.4mm	19.82mm
NRICP(V)	1.2mm	19.45mm

(c) Conventional global registration error

Figure 2.6: Traditional registration metrics shown in fig. 2.6c, fig. 2.6a and fig. 2.6b do not show the variation of registration performance with area which was less informative than SIRGn for evaluating quality of registration around the region of interest in the video clip.

were the largest input we used for 6 seconds of video. SIRGn runtime varied from 5 to 12 minutes depending on $ALG(V,G)$ input size. An example implementation is available under the BSD license at *Omitted for blind review*.

Traditional Metrics: Traditional metrics were computed frame to frame as they are not designed for global registration. Completely global RMSE and TRE metrics (fig. 2.6c) cannot represent time and it was not possible to use them in assessing how stable either $ICP(V)$ or $GEO(V)$ might have been. In addition registration drift could not be accounted for which led to misleadingly low error reports. Even when a time dimension is added to RMSE as in fig. 2.6a it was not possible to isolate the respiratory ROI.

Although we could use our G and C points to define mean TRE per frame as shown in fig. 2.6b, it was still not trivial to isolate the ROI. In contrast, the SIRGn overlays in figs. 2.6b and 2.6c clearly show regional variations of registration quality between $ICP(V)$ and $GEO(V)$ on the respiratory region, which was colored yellow for $ICP(V)$. SIRGn was most reflective of the more rigid registration $ICP(V)$ implemented with respect to $GEO(V)$ of all the evaluated metrics. In addition the ε threshold of SIRGn allowed us to see that the GEO algorithm was likely unsuitable for an abdominal AR guidance application requiring sub 5mm accuracy. This is in contrast to our evaluations with traditional metrics of RMSE and TRE where it is difficult to know if sub 5mm accuracy is possible in regions of interest.

2.4 Conclusion

We describe SIRGn – a novel metric for evaluating the quality of nonrigid registration algorithms on surgical 3D video. SIRGn indicates the quality of the registration for different parts of the 3D model by generating an overlay mesh that highlights areas where the registration approach succeeds and fails. This can serve as a way to evaluate different registration algorithms (both new and old) in various surgical scenarios. The ultimate goal is to provide a metric by which different registration algorithms crucial to enabling surgical augmented reality can be compared and evaluated. Our experiments demonstrate how SIRGn can be used to compare different registration methods. SIRGn evaluates a given registration algorithm by creating a mesh of registration quality over a 3D surgical video sequence using known good landmarks and our novel warp variance interpolation. We demonstrate how a SIRGn mesh can show relative registration quality between algorithms under evaluation, and also how an error tolerance can be specified when generating a registration quality mesh to highlight regions of unacceptable registration accuracy for a given application.

This chapter, in full, is a reprint of the material as it appears in the 14th International Symposium on Visual Computing (ISVC), Barrow, Michael; Ho, Nelson; Althoff, Alric; Tueller, Peter, and Kastner, Ryan, 2019. There are small changes in format and phrasing as a chapter within this larger paper. The dissertation author was the primary investigator and author of this paper.

Chapter 3

Surface Registration Accelerator

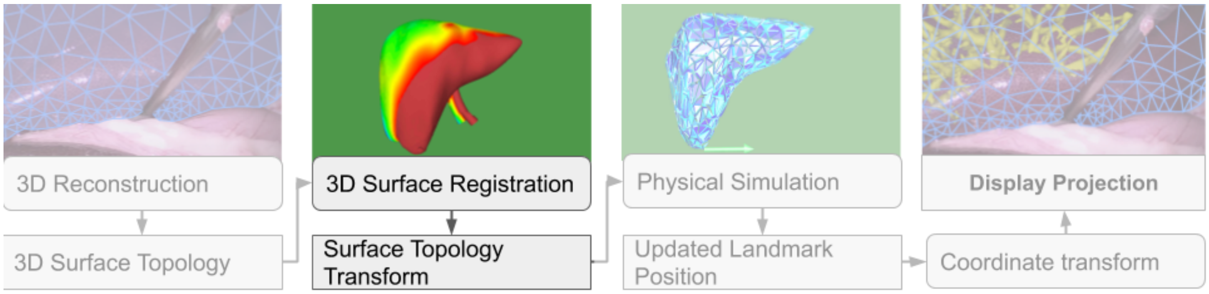


Figure 3.1: Our TRWL-S registration accelerator improves efficiency of the online 3D Surface Registration step in a CAIG pipeline.

Recall the four major online steps required for accurate CAIG outlined in chapter 1. In this chapter we describe “TRWL-S”, a method that applies to the 3D surface registration step. Figure 3.1 highlights the registration step in the CAIG pipeline to illustrate where our contribution is useful within a larger CAIG system. TRWL-S improves the speed of a registration algorithm whilst maintaining the accuracy and reducing power consumption. It therefore improves efficiency of CAIG.

Image registration is the fundamental computer vision problem that matches two or more scans of an object to each other, i.e., given two scans $X,Y \subset \mathbb{R}^3$, registration seeks the mapping $f : Y \mapsto X$ that corresponds to the smallest transform T where: $X = T(Y)$. In other words, registration indicates the simplest way to manipulate Y so the scans share the same coordinate system.

Registration has many real time applications such as SLAM, Stereo reconstruction, object tracking etc. [167, 4, 168]. A typical real time registration involves processing point cloud data from an input source (e.g., an RGBD camera) tens of times per second and mapping the 3D points from one frame to the next. We focus specifically on the 3D registration mapping function $f : Y \mapsto X$ in this work, which is a system bottleneck for real time performance in that it must perform a global optimization on the input data, and is a common step in all registration based applications.

Non-rigid 3D registration is a challenging but important class of registration where objects in a scan can deform non-rigidly. Fig 3.2 provides an example of a hand that deforms as the index finger flexes. Because the bones curl around each other, a non-rigid registration mapping function must be complex enough to consider all bones when registering an entire hand. As objects become less rigid, registration becomes more difficult as the number of articulation points increases.

(a) Hand scan X, fingers in open position

(b) Hand scan Y, index finger flexed

Figure 3.2: Non-rigid registration of a flexing finger. Our mapping strategy is to use the invariant "geodesic distances" in both X and Y to correctly label Y's features such that they map to the same features in X

Non-rigid registration is computationally hard because a non-linear optimization procedure is required to find the correct pairing of the features in Fig 3.2. In contrast, rigid image registration does not need to find a unique pairing of features since a linear decomposition on any sub-set of features can reveal the one mapping that is valid for all points in a scan. This makes real time non-rigid registration much more challenging. It is extremely usefull to solve difficult non-rigid registration problems in many applications [169, 170, 171], but unfortunately this has traditionally been hampered by the complexity of non-rigid mapping functions that can only achieve offline performance. For example, surgeons have long sought to use real time registration of soft tissue for image guidance applications [143, 136, 139, 142]. Frame rates must be above 10Hz to be good enough for image guidance and ideally closer to 25Hz [44].

The contributions of this work are: 1) A *Real Time Non-rigid Registration Accelerator* which demonstrates performance of 20 scans per second on a difficult medical registration problem. A distinguishing feature of the accelerator is that it does not require "shape templates" or "object priors" and is therefore suitable for difficult and simple non-rigid registration problems without re-engineering. 2) The *"TRWL-S" Algorithm.* We describe several novel methods we used to transform a state of the art offline memory bound algorithm into an online compute bound one. 3) *Energy Efficient Architecture.* We analyze performance of our new algorithm on our architecture and find a speedup of 600X with an 81X reduction in system energy.

The paper is organized as follows: **Section** 3.1 summarizes the state of the art in online non-rigid registration. **Section** 3.2 details how our algorithm enables our accelerator to have higher performance than a state of the art baseline algorithm. **Section** 3.3 describes the implementation of our novel registration architecture. **Section** 3.4 Evaluates the performance of the architecture. **Section** 3.5 Provides our conclusions on this work.

3.1 Related Work

Non-rigid 3D registration is computationally intensive. To the best of our knowledge there is no system capable of solving the general problem in real time. Instead, systems achieve fast runtime with trade offs to simplify their problem.

For example, much work has focused on real time pose estimation where object priors simplify the registration problem. Object priors are a set of rules describing plausible configurations for a canonical object. Registration requires finding a transform of a scan that best fit these rules and only works on objects that adhere to these rules. This is a simpler problem than performing registation with unknown content. The most common priors are object skeletons which can be articulated into the poses of interest [172, 173, 174, 175]. These approaches cannot be applied to more complex registrations, e.g., they are inadequate for registering pliable objects that deform without points of articulation.

Another approach possible is to define a registration template from the data stream itself. Newcombe et al. [167] avoid pre-defined priors by taking one image scan to be a canonical frame. The work achieves real time performance by regularizing (linearizing) the fitting of input scans to the canonical frame. They demonstrate that more complex real time registrations are possible with this approach as compared with skeleton templates; however the linear optimizer overconstrains the shape matching problem making registration of very flexible objects difficult.

Dou et al. [170] propose another real time system that does not require templates and is more robust than Newcombe's to large deformations. The registration is done using a deformation graph [176]. Although the results are compelling, a major drawback is that the deformation graphs must be defined by the user, which limits the generality of the technique.

In contrast, global optimization registration methods are more general since they do not require any object priors. For example, Chen et al. [162] achieved accurate results by formulating the optimization problem as a Markov Random Field (MRF). Each entry in the tensor is a cost associated with pairing geodesic distances between features. While the quality is high, the optimizer run time is prohibitively slow for real time applications.

Real time global non-rigid registration is therefore possible if an optimizer exploring the MRF problem can do so in real time. In addition if the number of features "n" that can be described by the MRF is large enough, this approach would be suitable for difficult registration problems.

There have been recent advances in GPU, FPGA, and ASIC acceleration for efficient message passing based global optimizers which explore MRF's in linear time [2, 3, 4, 5]. Unfortunately these approaches are revealed to be memory bound as they must stream the MRF into the accelerated optimizer. Fig 3.3 illustrates accelerator memory bound for MRF problems sizes of Chen's approach[1]. Fig 3.3 shows it is not possible to solve MRF registration problems in real time for $n \geq 60$ with existing accelerators because the MRF cannot be loaded into the accelerator quickly enough.

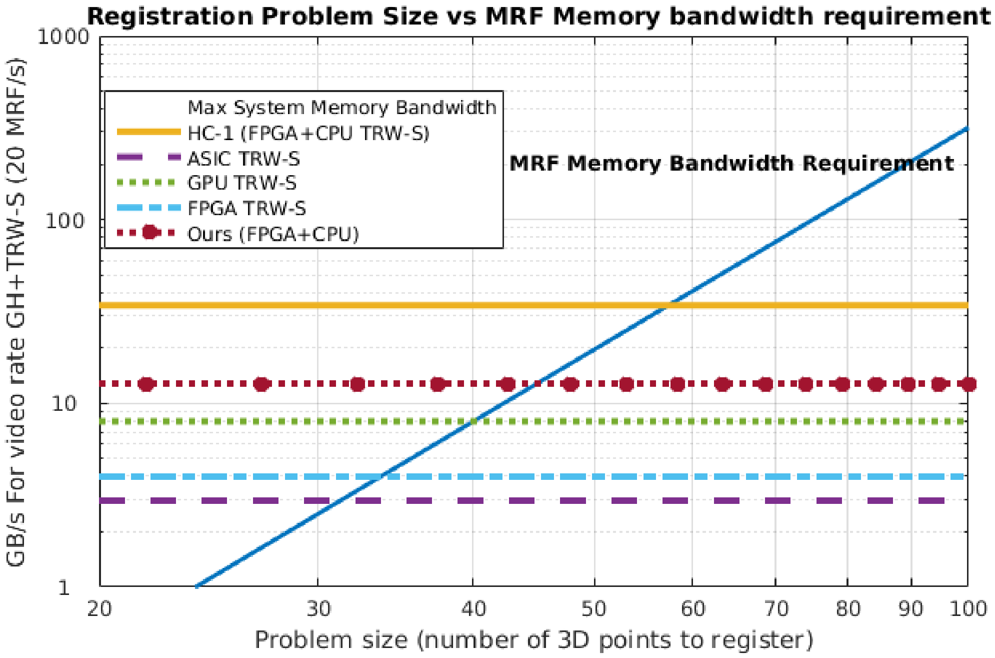


Figure 3.3: MRF memory bandwidth scaling with problem size (n) [1]. No proposed global optimizer accelerator system (horizontal trends) is suitable for difficult non-rigid registration since the MRF must be streamed in. The memory bandwidth available for transferring a global registration problem MRF to the proposed accelerators [2, 3, 4, 5] is not sufficient for a moderately complex registration problem of n=60 feature points.

3.2 Registration Algorithm

Our Tree Reweighted Loaf Slice (TRWL-S) algorithm builds on a global non-rigid registration baseline described by Chen et al. [162] which registers scans using geodesic distances (see Fig 3.2). Fig 3.4 illustrates how these invariant distances can be summed into a $n \times n$ matrix that gives feature to feature distances between all n features and describes a global relationship between all the features in a scan.

	f0	f1	f2	f3
f0	0	1	2	3
f1	1	0	3	4
n f2	2	3	0	5
f3	3	4	5	0

(a) Scan $\mathbf{X} \in \mathbb{R}^{n \times n}$ feature to feature Geodesic distances

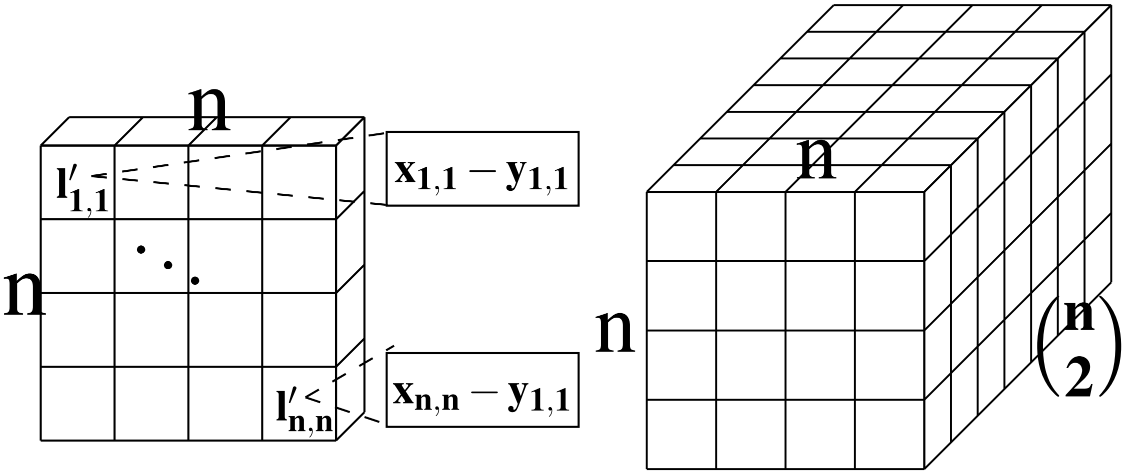
	f?	f?	f?	f?
f?	0	3	2	1
f?	3	0	5	4
n f?	2	5	0	3
f?	1	4	3	0

(b) Scan $\mathbf{Y} \in \mathbb{R}^{n \times n}$ feature to feature Geodesic distances

Figure 3.4: Matrix 3.4a represents a global intrinsic model of the hand from Fig 3.2. The model is a list of geodesic distances between every pair of features; $x_{i,j} = |f_i - f_j|$ where: $i, j =< 1, \ldots, n >$. Matrix 3.4b is the geodesic distances between features as seen in Scan Y. The goal is to match the features from Scan Y to X using the known distance mappings in Matrix 3.4a, i.e., find the correct permutation of i, j for Matrix 3.4a to Matrix 3.4b s.t. $x_{i,j} = y_{i',j'} = |f_i - f_j|$

To register **X** and **Y** in Fig 3.4, the baseline constructs a MRF by applying a geodesic distance heuristic (GH) to each combination of **X** and **Y**, and then uses a Tree Reweighted Message Passing (TRW-S) global optimizer to find a best matching. Since GH+TRW-S : $Y \mapsto X$, we refer to the baseline as "GH+TRW-S" to denote the registration function we accelerate. Although GH+TRW-S is global and requires no object specific priors, the trade off is it has greatly increased computation compared with the related work and online performance is very difficult to achieve.

The first compute heavy step is: *Geodesic Heuristic (**GH**)* where the registration is mapped as an energy minimization MRF problem by calculating a energy cost of matching pairs of points in X and Y. This energy cost function is a heuristic form of geodesic distance matching where a filtering and weighting scheme is applied to the geodesic distances to make the registration more robust to imperfections of 3D scanning systems [162]. The GH algorithm results in an $O(n^4)$ sized MRF which is illustrated in Fig 3.5.



(a) A MRF "Slice" (L') is the n^2 elements that make $\mathbf{X} - y_{i,j}$

(b) The complete MRF "Loaf" (L) is all $\binom{n}{2}$ Slices

Figure 3.5: A MRF registration problem is constructed using two matrices X and Y of geodesic distances (see Fig 3.4). A good labeling is one where geodesic distances are the same ($\mathbf{X}_{a,b} - \mathbf{Y}_{i,j} = 0$). Therefore $\mathbf{X} - y_{i,j}$ are all n^2 potential matches for one $y \in Y$. This means checking all potential matches requires creating an MRF by subtracting all $y \in Y$ from X. We abstract this as a 3D tensor "loaf of bread": $L = [X - y_{1,1}, \ldots, X - y_{n,n}]$ of order $O(n^4)$.

The second step is a *Global Optimal Matching* search using (***TRW-S***): TRW-S uses a global convex non-linear energy optimization to the explore the MRF problem and find a optimal matching candidate solution in $O(kn^4)$ time [6], where the registration solution calculates n matches and each scan has n features [162]. TRW-S is a popular MRF solver due to its fast convergence and its ability to compute a lower bound energy. It can report a theoretical best solution (the lower bound) which helps to choose a "k" constant for the overall $O(kn^4)$ complexity [177]. Full details on GH+TRW-S can be found in [162] and [6].

The major challenge to accelerating this algorithm is efficiently constructing $O(n^4)$ MRFs during the GH step, and then rapidly exploring this $O(n^4)$ problem space in the TRW-S phase.

GH+TRW-S does not map well to most parallel architectures. In particular TRW-S has data dependencies that make it difficult to schedule.

TRW-S and other message passing based accelerators typically optimize DRAM streaming of the $\binom{n}{2}$ messages [2, 3, 4, 5]. However, we found GH+TRW-S to be memory bound when the MRF is ignored, since the TRW-S solver must stream in an MRF from the GH step as shown in Fig 3.3. The problem size does not need to be very large before the $O(n^4)$ MRF exceeds typical cache sizes and accelerators become bottlenecked. e.g., a 64 point registration requires 66 MB of cache [1].

3.2.1 Dynamic MRF Generation

Fig 3.3 shows that video rate performance is not possible using the baseline GH+TRW-S algorithm on any reviewed accelerator. 464GB/s bandwidth is required to achieve 20 scan/s video rate on a 110 node registration. Typically, hardware acceleration platforms have significantly less bandwidth. For example, our Arria 10 + Xeon hardware can only stream the MRF to a TRW-S accelerator at 12GB/s. To reduce this bandwidth requirement, our algorithm uses a novel approach called *dynamic Markov random field generation (DMRF).*

Our DMRF algorithm dynamically computes the "slices" shown in Fig 3.5a as they are needed by TRW-S. DMRF is able to reduce MRF storage space because the number of independent random variables is actually far lower than n^4 for registration. Fig 3.4 shows only $2n^2$ variables are used to construct the $O(n^4)$ MRF for **X** and **Y** depicted in Fig 3.5b. DMRF dynamically evaluates the GH energy cost function:

$$E = (\exp(-\min(x/\sigma, y/\sigma)) \cdot \min(|x-y|, \tau)) \cdot q \tag{3.1}$$

Equation (3.1) is evaluated for every "cell" in each "slice" of Fig 3.5b where; σ and τ are GH filter and weight constants respectively, and q is a normalization factor for numerical stability [162]. Dynamic evaluation means only the $2n^2$ independent MRF variables are transferred to our accelerator. Fig 3.6 shows our platform is compute bound for larger problems with this approach. eg. 0.08GB/s is used for 20 scan/s 110 node DMRF transfer versus 464GB/s for MRF.

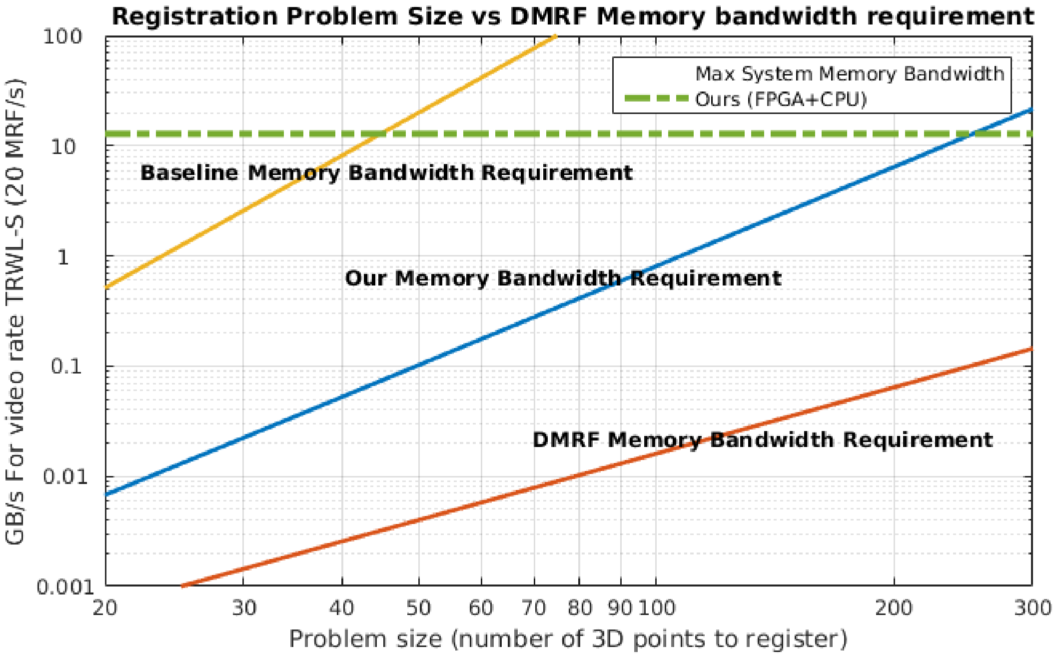


Figure 3.6: DMRF removes the MRF transfer bottleneck ("Baseline" trend) which changes the problem from memory to compute bound ("DMRF" trend). However, scaling on our platform is now constrained by the "**M**" optimizer messages of Alg 1 ("Our" trend) and this is discussed in Sec 3.4.4

Provided X and Y are available, Eqn 3.1 can be evaluated simultaneously for the entire MRF. This makes generating loaf slices with DMRF an effective approach. Our algorithm integrating our DMRF generation into TRW-S optimization is described in Alg 1. We call this modified algorithm "TRWL-S".

Algorithm 1 TRWL-S

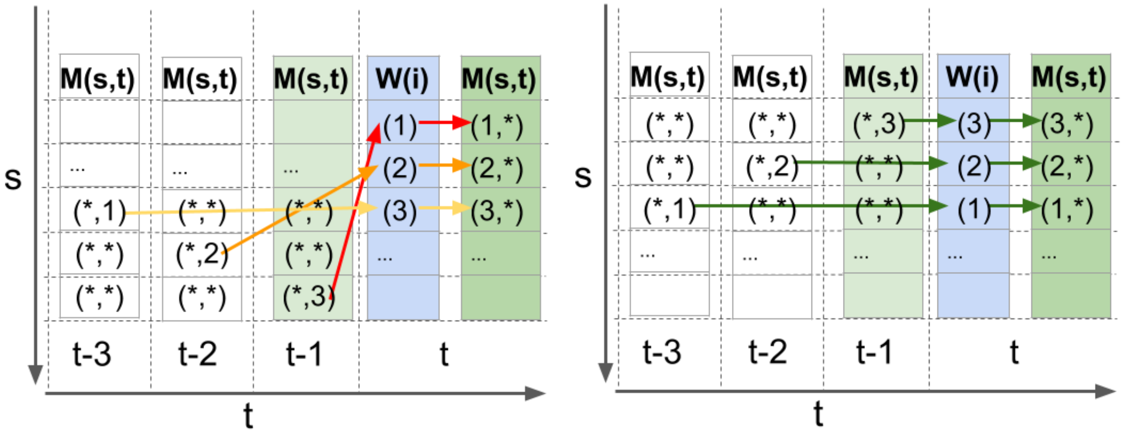
```
procedure DMRF($\mathbf{X}, \mathbf{Y}, s, t, q$)
    $y = \mathbf{Y}$[s,t]
    return $E(\mathbf{X}, y).q$                                  ▷ see Eqn 3.1
procedure UD($M, L', W, s, t$)
    off = $\Gamma[s,t] \circ W - M$                              ▷ ∘ is Hadamard product†
    $G'[1:n,*] = L'[1:n,*] +$ off
    $M'$ = COLMIN($G'$)
    return $M'$ - MIN($M'$)
procedure TRWL-S($\mathbf{U}, \mathbf{X}, \mathbf{Y}, q, j$)     ▷ $\mathbf{U}$ is optional$^*$
    $\mathbf{M}[*,*] \leftarrow [0]$
    while $j$-=1 $> 0$ do                                        ▷ j is number of passes
        $W_f \leftarrow W_b \leftarrow \mathbf{U}$
        for $s = n$; $s > 0$; $s$-=1 do                          ▷ Back pass (TRW-S)
            for $t$=$s$-1; t>0; $t$-=1 do
                $W_b[t]$+= $\mathbf{M}[t,s]$
                $W_b[t]$+= $\mathbf{M}[s,t]$
            for $t$=$s$-1; $t$>0; $t$-=1 do
                $L' \leftarrow$ DMRF($\mathbf{X}, \mathbf{Y}, s, t, q$)
                $\mathbf{M}[t,s]$= UD($\mathbf{M}[t,s], L', W_b[t], s, t$)$^\dagger$
        for $t$=$n$; $t$>0; $t$-=1 do                            ▷ Front pass (ours)
            for $s$=0; s<$t$-1; $s$+=1 do
                $W_f[t]$+= $\mathbf{M}[t,s]$
                $W_f[t]$+= $\mathbf{M}[s,t]$
            for $s$=0; s<$t$-1; $s$+=1 do
                $L' \leftarrow$ DMRF($\mathbf{X^T}, \mathbf{Y}, s, t, q$)
                $\mathbf{M}[s,t]$= UD($\mathbf{M}[s,t], L', W_f[t], s, t$)$^\dagger$
        $S \leftarrow$SLN($\mathbf{M}, \mathbf{G}, \mathbf{U}$)$^\dagger$
    return $S$
```

$*\mathbf{U}$ *is an optional* ***TRW-S*** *seed and may be set to zero* [6]

†Γ*, message update (UD) and solution (SLN) detailed in* [6]

3.2.2 Scheduling Optimizations

TRWL-S must be scheduled carefully due to the data dependency on $W[t]$ in the UD() step. Alg 1 shows two scheduling options for the two optimization passes of TRWL-S "Back Pass" uses the naive TRW-S scheduling which is sub-optimal. The serial dependency issue is illustrated graphically in Fig 3.7a. Our "Front Pass" schedule in Alg 1 ensures that the first $\mathbf{M}[t]$ required to be added to a $W[t]$ in the next loop of the front pass is computed first and is shown in Fig 3.7b. This scheduling optimization is important for effectively speeding up TRWL-S. Data dependencies prevent a fully parallel schedule and so pipelining is a better execution scheduling option. Without our schedule, TRWL-S would suffer pipeline stalls of up to $n \times (\text{UD()} + \text{MRF()})$ time. Since both of these functions are a $O(n^2)$ compute bottleneck, total stall time without our schedule is on the order of $O(n^3)$.



(a) Data dependencies using naieve TRW-S schedule in Alg 1

(b) Data dependencies using our schedule in Alg 1

Figure 3.7: TRWL-S serial data dependencies between **M** and **W**. s is the inner loop and is completed once before traversing along t. Fig 3.7a is the "Back Pass" TRW-S schedule in Alg 1 and Fig 3.7b is our "Front Pass" schedule in Alg 1. $\mathbf{M}[s,*]$ depends on $\mathbf{W}[s]$ from $t-1$ which in turn depends on $\mathbf{M}[*,s]$ from $t-2$. Our schedule in Fig 3.7b is better suited to pipelining as we schedule the compute of **M** to meet the dependencies of **W** at $t+1$ as early as possible.

3.2.3 MRF Optimization

The MRF cost function in Eqn 3.1 uses the *exp* function which is expensive in terms of both resource usage and latency. To improve this, we define a transformed cost function in Eqn 3.2:

$$E = \max(a,b) \cdot \min(|x-y|,\tau) \tag{3.2}$$

where $a = qe^{-x/\sigma}$ and $b = qe^{-y/\sigma}$. The transformation is valid because exp is a monotonic, increasing function of its argument, and $\max(-u,-v) = -\min(u,v)$ is an identity. The values of a and b are pre-computed on CPU using $O(n^2)$ work. We include the (positive) scaling term q in a and b when factoring out the expensive exp to reduce even further the total compute required for the $O(n^4)$ FPGA work .

3.2.4 Precision Optimization

In order to maximize the performance, we aim to saturate the FPGA multiplier blocks and other arithmetic resources. Key to this is using fixed point arithmetic since the target Arria 10 FPGA has more integer DSP multipliers than floating point multiplier blocks.

The GH+TRW-S baseline uses 64 bit floating point numbers. The Arria 10 DSP multipliers are 18 bits wide, so ideally we would prefer to reduce baseline precision to match the DSP multipliers. At the same time, in order to maintain accurate registration, the accelerator can not arbitrarily reduce the TRWL-S word width.

We analyzed the impact of reduced word width by sweeping fixed point precision of GH+TRW-S on our data sets to verify reduced precision did not negatively impact registration performance. Fig 3.8 shows that the energy minimization (registration capability) of GH+TRW-S is not greatly affected until we reduce the data representation to a 8 bit fixed point. We conservatively choose 16 bit fixed point since this precision maps well to the Arria 10 DSP multipliers, reduces memory footprint by a factor of 4X over GH+TRW-S, and does not greatly impact the registration quality on our tested data sets. we do note that other data sets may vary although 16 bits is acceptable across our two tested data sets.

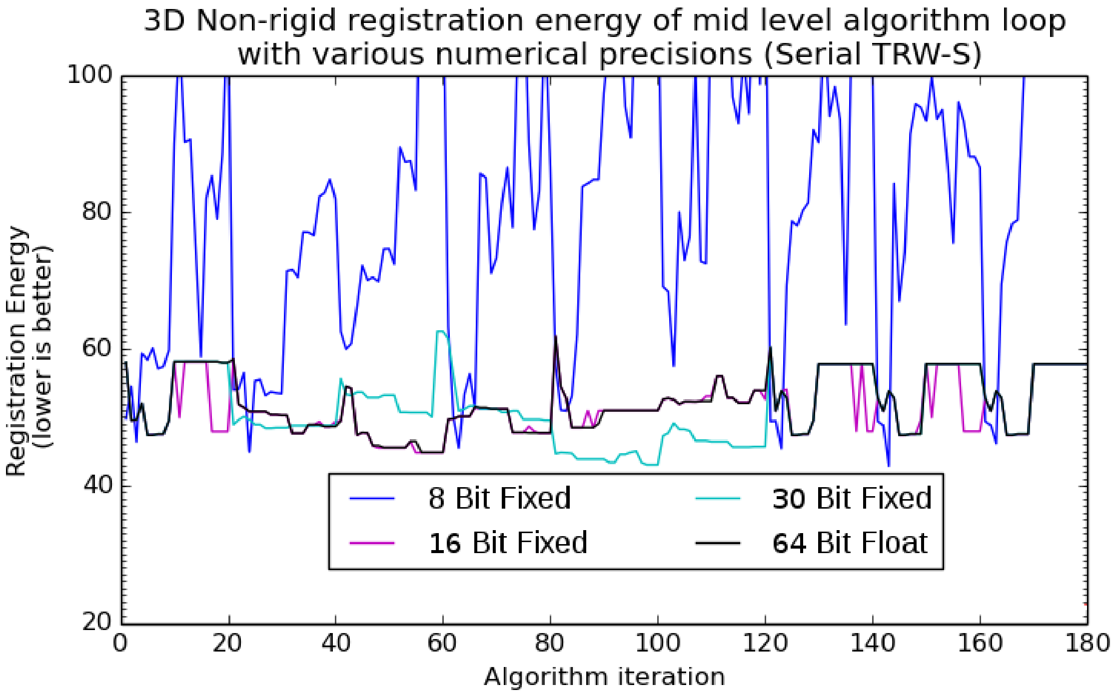


Figure 3.8: Registration energy vs. the algorithm iteration for different data types. A lower energy indicates a more accurate registration. The 64 bit float, 30 bit fixed, and 16 bit fixed all have similar registration energy. The 8 bit fixed is significantly worse.

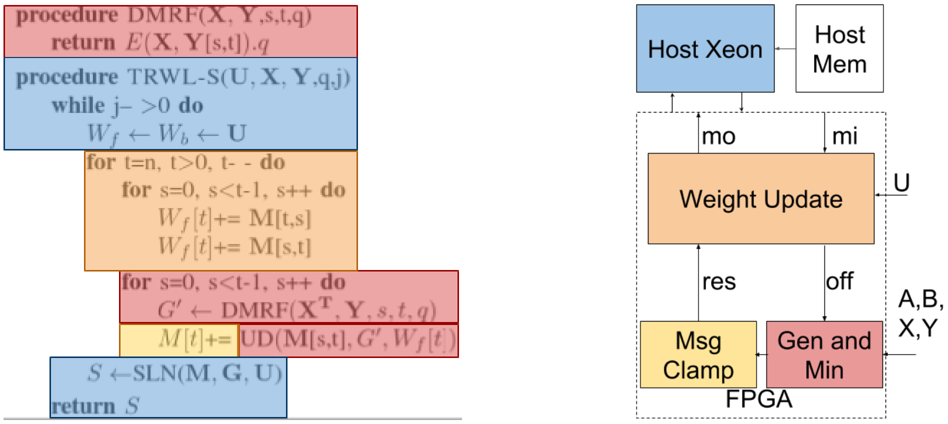
3.3 Accelerator Architecture

3.3.1 System Overview

Our accelerator is built using a 14 core Intel Xeon Broadwell coupled to a Altera Arria 10 FPGA via QPI and PCIe interconnects. The portions mapped to the Xeon were implemented in C++. The portions mapped to the FPGA were designed using Chisel3 [178]. The CPU to FPGA memory interface was built using Intel's *Rapid Design Methods for Developing Hardware*

Accelerators methodology [179], which is optimized for CPU to FPGA data transfers of cache line granularity, i.e., 32 16 bit fixed precision numbers per FPGA cycle. Because of the CPU/FPGA memory interface, it is efficient to process data in vectors of 32 numbers and all block to block data transfers in Fig 3.9b use buffers 32 elements wide.

Fig 3.9 provides a high level block diagram of our heterogeneous non-rigid 3D registration accelerator. The Host Xeon provides high-level control, executes initialization procedures, and performs the mathematical transforms needed for registration. The functions DMRF() and UD() from Alg 1 are mapped to the FPGA since they can benefit from a parallel implementation. **W** operations from Alg 1 are also mapped to the FPGA. This is because the accelerator would suffer execution starvation from serial dependencies on **W** (see Fig 3.7) if the **W** variables had to be maintained by the CPU and streamed in as needed. The blocks in Fig 3.9b are each detailed in Sections 3.3.2 to 3.3.5.



```
procedure DMRF($\mathbf{X}$, $\mathbf{Y}$,s,t,q)
    return $E(\mathbf{X}, \mathbf{Y}[s,t]).q$
procedure TRWL-S($\mathbf{U}$, $\mathbf{X}$, $\mathbf{Y}$,q,j)
    while j– >0 do
        $W_f \leftarrow W_b \leftarrow \mathbf{U}$
        for t=n, t>0, t- - do
            for s=0, s<t-1, s++ do
                $W_f[t]$+= M[t,s]
                $W_f[t]$+= M[s,t]
            for s=0, s<t-1, s++ do
                $G' \leftarrow$ DMRF($\mathbf{X}^{\mathbf{T}}$, $\mathbf{Y}$, $s, t, q$)
                $M[t]$+= UD(M[s,t], $G'$, $W_f[t]$)
        $S \leftarrow$SLN(M, $\mathbf{G}$, U)
    return $S$
```

(a) TRWL-S pass with hardware mapping highlighted

(b) Xeon/Arria 10 TRWL-S accelerator block diagram

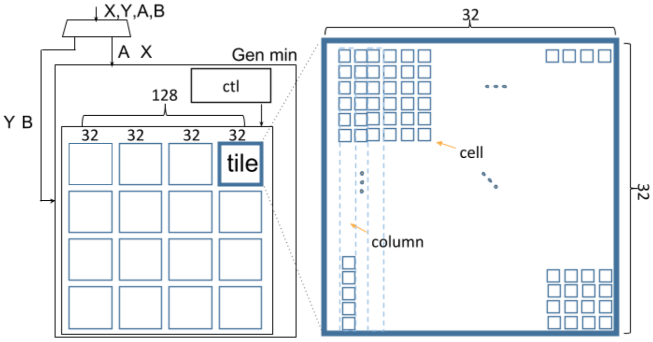
Figure 3.9: Heterogeneous TRWL-S accelerator block diagram. Fig 3.9a colours a TRWL-S pass to indicate how we partition the algorithm on our accelerator. Fig 3.9b is our Xeon/Arria 10 3D non-rigid registration accelerator. Each hardware block is coloured to match the portion of TRWL-S described in Alg 1 it is responsible for. We omit the Backward pass for brevity.

3.3.2 Gen and Min

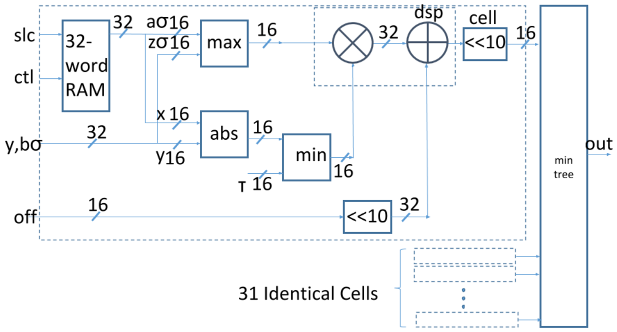
"Gen and Min" in Fig 3.9b implements DMRF() and the COLMIN() of UD() in Alg 1 since these are the compute bottleneck of TRWL-S, both having $O(n^2)$ complexity per slice. Both DMRF() and COLMIN() are SIMD which our accelerator takes advantage of as shown in Fig 3.10. The block depicted in Fig 3.10a shows how larger matrices, for example, 128×128, are mapped to 4×4 virtual tiles using time-division multiplexing. Our architecture physically implements 32×32 compute cells. The tile arrangement can be modified at runtime to describe registration accelerators for different problem sizes (multiples of 32.) Eqn 3.2 shows that **A** and **X** are constant factors for every loaf slice and so our architecture loads the appropriate a and x into the cell ram shown in Fig 3.10b once only per registration. We use double buffering to allow tiles to begin work as **A** and **X** are loaded from CPU. Each cell implements Eqn 3.2 for DMRF() and adds "off" as the second step of UD() in Alg 1. Each cell's output is then fed to COLMIN() hardware which is shown as the min tree block in Fig 3.10b. Min tree performs the n wide vector min using a log tree arrangement with 32 word inputs. Processing a slice takes 16 cycles (one cycle for each tile.) The gen and min unit is deeply pipelined (14 stages) giving a L' slice per FPGA clock ratio of 1:1.

3.3.3 Message Clamp

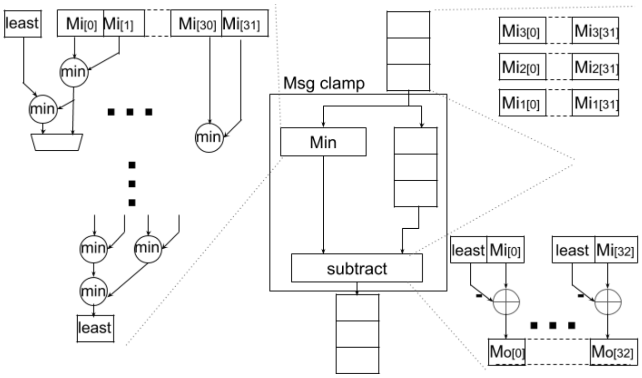
Msg Clamp or "message clamp" in Fig 3.10c implements the final MIN() step of UD() in Alg 1. Using the separate message clamp block simplifies our architecture. It is broken out from "gen and min" (Section 3.3.2) because MIN() is an aggregating step that is not a compute bottleneck and is parallel in an orthogonal way to COLMIN() of the gen and min block. Msg clamp uses the same log tree vector min concept as gen and min but cycles the $least$ output several times through the tree. This is because the input from gen and min is a cache line (32 numbers) wide and so *least*= MIN(x) should be run once for every tile of gen and min. The FIFO in the



(a) Gen and Min tile and cell architecture



(b) Cell and colmin block



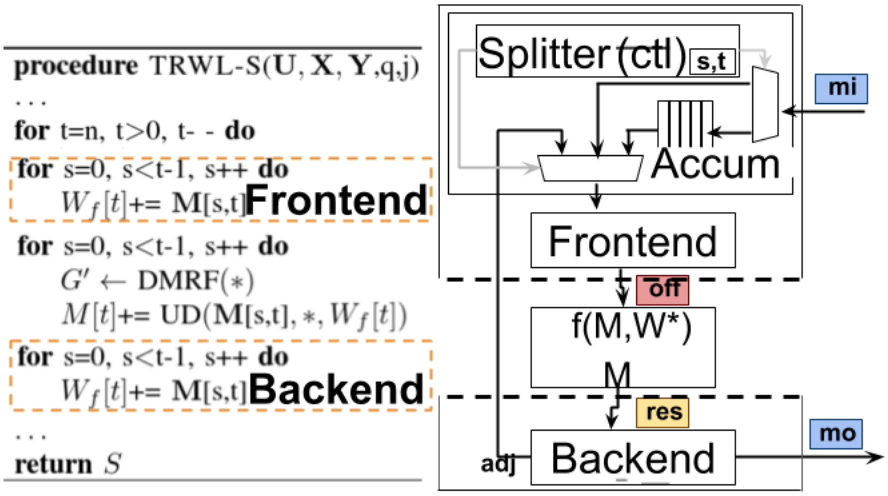
(c) Msg Clamp block

Figure 3.10: Gen and Min block. This block is dedicated to the $O(n^2)$ per slice compute bottleneck of TRWL-S. It computes DMRF() loaf slices from Alg 1 and n vector mins of that slice. Vector mins are the bottleneck of UD() from Alg 1. We achieve high performance with an array of 32×32 cells each containing a 14 stage pipeline as shown in Fig 3.10a and detailed in Fig 3.10b.

clamper delays arrival of results from the first tile to a 32 wide vector subtract until a true MIN() has been computed. **M**[.,.] is computed using vector subtract and streamed to “Weight Update”

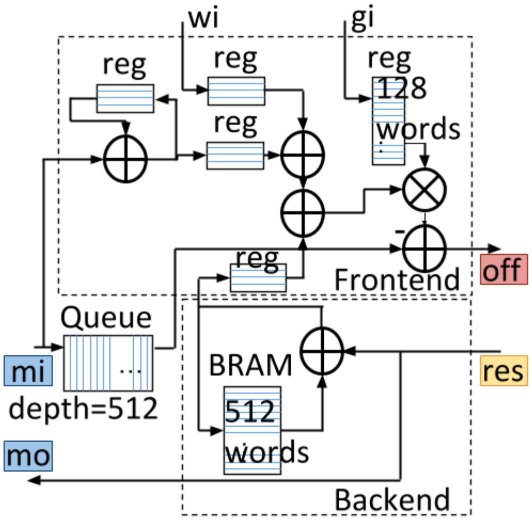
3.3.4 Weight Update

“Weight Update” in Fig 3.11 maintains the **W** weight variables in Alg 1 because they are on the critical path to meet the serial dependencies of TRWL-S and round trip latency is too high to manage **W** on CPU.



```
procedure TRWL-S(U, X, Y,q,j)
...
for t=n, t>0, t- - do
    for s=0, s<t-1, s++ do
        $W_f[t]$+= M[s,t]
    for s=0, s<t-1, s++ do
        $G' \leftarrow$ DMRF($*$)
        $M[t]$+= UD(M[s,t], $*$, $W_f[t]$)
    for s=0, s<t-1, s++ do
        $W_f[t]$+= M[s,t]
...
return $S$
```

(a) Partitioning and scheduling scheme of the Weight Update block



(b) Data flow details of the Weight Update block

Figure 3.11: Control and Data flow views of the Weight Update Block. Fig 3.11a shows a rescheduling of **W** computation that allows dependency forwarding of Alg 1 from t to $t+1$. Fig 3.11b details **W** processing. We use a local copy of **W** to efficiently compute dependencies. We perform canonical tree re-weighting with Γ in the "frontend" only for efficiency [6].

Rather than suffer the round trip penalty, **W** is stored on the FPGA in BRAM and is optionally initialized to **U** by the CPU. We re-order weight update as shown in Fig 3.11a so that the dependency of $t+1$ can be met by forwarding $\mathbf{M}[.,x]+\mathbf{W}[x]$. The forwarding stage is shown as the "Backend" in Fig 3.11a. We manage meeting the **W** dependencies of gen min (Sec 3.3.2) with the "Frontend" block in Fig 3.11a. This block tracks s and t, using these variables to route the correct message, $\mathbf{M}[.,.]$ and weight $\mathbf{W}[.]$. If the required dependency can't be satisfied, weight update will stall the accelerator. In practice stalls are negligible (0.2% stall cycles per MRF). If dependencies are met, $\mathbf{W}[.]$ and $\mathbf{M}[.,.]$ are scaled by Γ as shown in Fig 3.11b which is the first step of Alg 1 UD(). The result "off" is then passed to "gen and min".

3.3.5 CPU

Our 3D registration accelerator is heterogeneous since both CPU and FPGA are used in the data path. In addition to data processing, the CPU is also responsible for control. On the control path; The CPU starts TRWL-S and determines the number of iterations. We hard coded j in Alg 1 to be 20 iterations to match the GH+TRW-S base line for performance comparison. However, termination can also be determined dynamically by comparing the SLN() registration energy (solution quality) to either the lower bound [6] or an experimentally determined value.

The CPU can also optionally specify registration candidate solutions at each pass via the **U** parameter. We use the same **U** as the baseline for our performance evaluation.

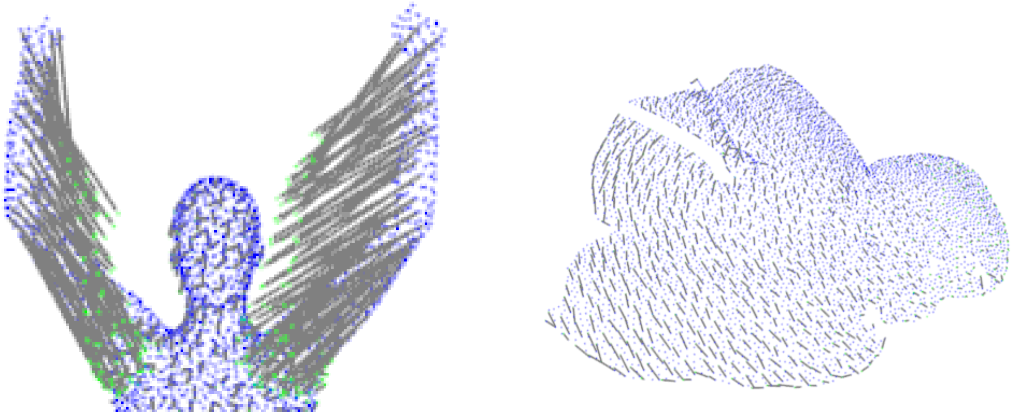
On the data path; The CPU pre-processes registration problems by computing **A** and **B** before starting TRWL-S as described in Section 3.2.3. Other pre-processing steps include computing the baseline q normalization factor [162] and converting all numbers to 16bit Fixed point representation. Since the **M** messages require $O(n^3)$ storage, the CPU streams them off and on to the FPGA as needed. Finally the CPU decodes the 3D registration solution by implementing SLN() [6].

3.4 Evaluation

We evaluate our system accelerator performance against a software implementation of the baseline GH+TRW-S and software TRWL-S. We compare only TRWL-S : $Y \mapsto X$ against GH+TRW-S : $Y \mapsto X$ since we do not re-implement the input stage (geodesic distance computation) or the output stage (expansion move upscaling). The baseline uses a state of the art implementation of TRW-S which has been enhanced with openMP to run a bottlenecking function (UD()) from Alg 1 in parallel on multiple CPU hardware threads and with SIMD CPU instructions when possible [1]. Our TRWL-S accelerator significantly outperforms the single threaded baseline implementation ($\approx$ 600X faster). This would make a holistic baseline software to accelerator comparison uninteresting. Therefore in this section we use the baseline to evaluate registration quality of the accelerator, but use a software TRWL-S implementation to compare other metrics. Specifically we evaluate accelerator run time performance and system power consumption of the accelerator against software TRWL-S. Finally we evaluate the scalability limitations of our TRWL-S non-rigid 3D registration accelerator.

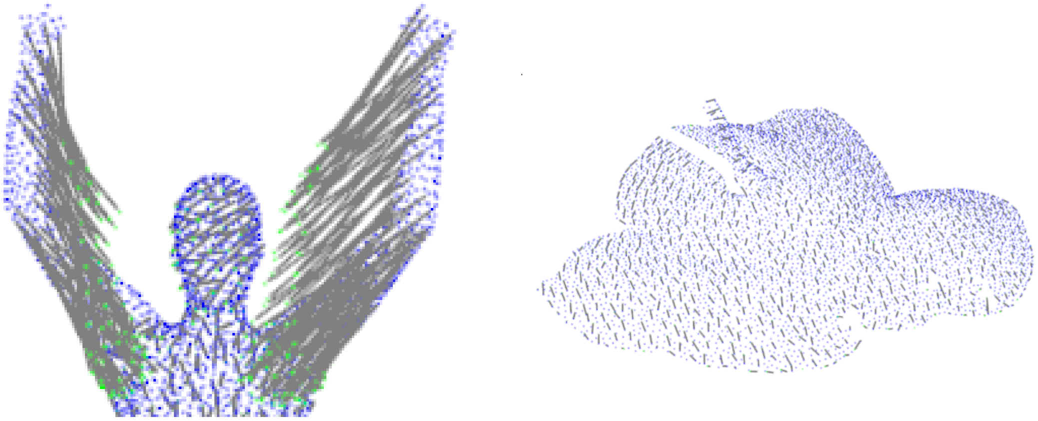
3.4.1 Registration Quality

We compared registration performance of TRWL-S to the baseline using the FAUST pose estimation database [180] and our own surgical video data set. Figure 3.12 shows the results. We calculate the mean difference of solution energy (registration quality) between baseline TRW-S and TRWL-S in table 3.1 across our data sets. The average energy difference is small since the algorithms are equivalent except for the reduced numerical precision and registration results are therefore very similar. Figures 3.12a to 3.12d show a qualitative comparison of the 3D registrations of both algorithms on the two different data sets. In general TRWL-S appears to have good agreement with the baseline.



(a) Pose estimation: baseline used for 3D registration

(b) Liver surgery: baseline used for 3D registration



(c) Pose estimation: our accelerator used for 3D registration

(d) Liver surgery: our accelerator used for 3D registration

Figure 3.12: 3D nonrigid registration results of TRWL-S (Fig 3.12c, 3.12d) and the baseline (Fig 3.12a, 3.12b)

Table 3.1: Relative comparison between TRWL-S and GH+TRW-S registration quality. δ is the mean energy (quality) difference for a set of similar registrations. A small δ is desirable and indicates similar TRWL-S and GH+TRW-S registration results. *n* is the problem size and “Scans” are the number of registrations used to measure δ.

Dataset			Mean δ
Name	Scans	n	(lower is better)
Liver Surgery	25	96	1.3%
FAUST	50	96	1.9%
Liver Surgery	25	128	1.8%
FAUST	50	128	1.7%

3.4.2 Accelerator Performance

We implemented the accelerator on a Arria 10 10A115U3F45E2SGE3 FPGA integrated into a Xeon-FPGA "HARP" platform. We used Chisel3 [178] and the *Rapid Design Methods For Developing Hardware Accelerators* [179] methodology for implementing the system. Chisel3 enables highly parameterized, clock accurate RTL descriptions to be developed in a concise and effective manner. We set the active matrix size to be 32×32 so that the *Gen and Min* component uses 1024 of the 1518 DSP blocks available on the FPGA. All components of the system together occupy 42% of the available ALMs (this includes both the Blue Stream, the interface to the Xeon provided by the platform, and the Green Stream, the hardware specific to this accelerator.) Timing converged to 400 MHz in the Blue Stream and 100 MHz in the Green Stream (no timing violations.) Other FPGAs with more compute resources (e.g., the Stratix 10 with up to 5760 DSP blocks) and higher clock rates (200 MHz or more can likely be achieved with more implementation work) would improve the performance proportionally.

3.4.3 System Power Consumption

We measured power consumption on the physical platform. The measurement system provides a breakdown by: CPU power, DRAM power, Core FPGA power, and Other FPGA power. For the software-only case, we measured power on the same system. Removing the FPGA components from the software-only case, we see a 69.6 W for software only vs. 72.8 W for the accelerated system. The runtime for the accelerated system is $\approx 84\times$ smaller (compared with a single core software-only implementation), resulting in an energy per computation benefit of $\approx 81\times$. The runtime and energy benefits are less when compared to a multi-threaded software implementation ($10\times$ and $15\times$, respectively) but still significant.

Table 3.2: Measured Power for the accelerator computation. "FPGA Accel" are power metrics for our non-rigid 3D registration TRWL-S accelerator system. "SW (1 third)/(14 thirds)" are metrics for software-only TRWL-S registration using 1 and 14 CPU hardware threads respectively. Adding threads makes a CPU compare increasingly favorably with FPGA, however the FPGA always maintains a large energy and speed advantage in our system.

	CPU (W)		FPGA (W)		Perf per Frame	
Name	Core	DRAM	Core	Other	time(s)	energy(J)
FPGA Accel	24.2	28.1	8.7	11.8	0.05	3.6
SW (1 third)	41.5	28.1	(7.1)	(11.8)	4.22	294.0
SW (14 thirds)	86.0	28.1	(7.1)	(11.8)	0.50	55.2
Ratio (1)	1.71	1.00	0	0	84.5	80.8
Ratio (14)	3.55	1.00	0	0	10.0	15.2

3.4.4 Scalability

Figure 3.13 shows both memory bandwidth and compute bandwidth limitations of TRWL-S on the same log-log graph.

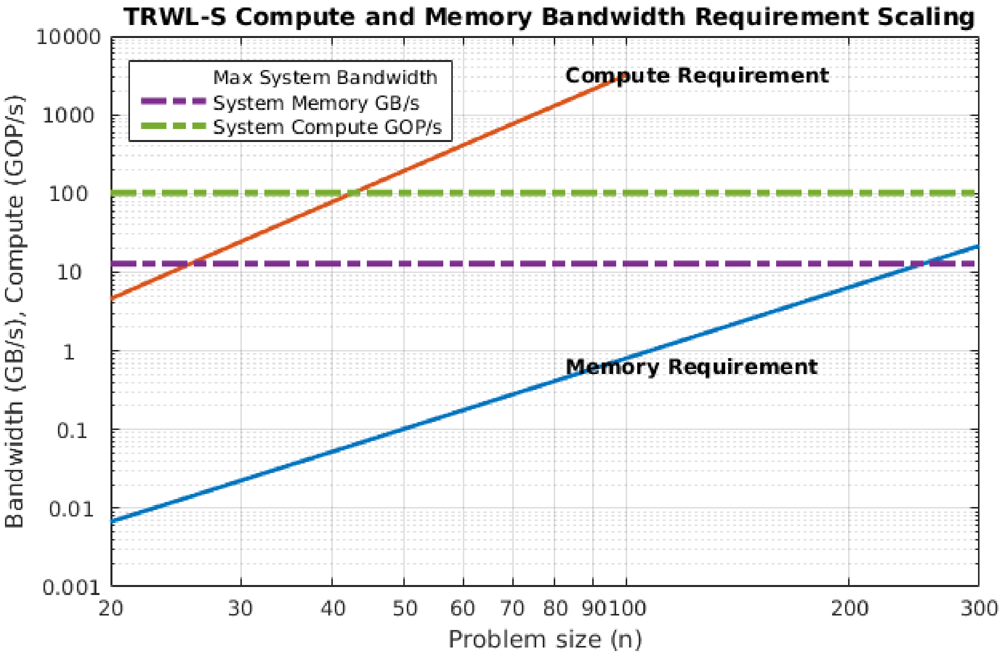


Figure 3.13: TRWL-S scaling on a Xeon Aria 10 platform. Memory bandwidth scaling is dominated by message streaming which is $O(n^3)$. Compute scaling is dominated by the $O(n^4)$ DMRF computation. Compute must improve by 17x (using a faster clock or larger FPGA) before platform bandwidth is the performance limiter (at n=256).

The limitations of our TRWL-S algorithm are: 1) the $O(n^3)$ bandwidth requirement for transferring $\binom{n}{2}$ length n messages to and from memory, and 2) the $O(n^4)$ compute required to perform the *Gen and Min* operations.

Compute can be increased by clocking the system faster or utilizing an FPGA with more resources (DSP blocks). The horizontal line showing the compute limit would increase proportionally. It can increase 17x before the platform bandwidth becomes the bottleneck around $n = 256$.

3.5 Conclusions

We describe a heterogeneous CPU/FPGA accelerator for real-time non-rigid 3D registration. The design uses a MRF transform and scheduling optimizations to achieve a clinically acceptable 20 registrations per second performance. Experimental results show clear performance benefits of the accelerator. Our system achieves $\approx 600\times$ speed up with a maximum 1.9% difference in registration quality over a software only TRW-S non-rigid 3D registration baseline. Additionally we find a $\approx 84\times$ speed improvement and $\approx 81\times$ energy reduction of our heterogeneous TRWL-S architecture versus a software only implementation. In future work we plan to further reduce the 1.9% quality difference by tuning the registration parameters of TRWL-S and to investigate current scaling limitations with our approach.

This chapter, in full, is a reprint of the material as it appears in the 28th International Conference on Field Programmable Logic and Applications (FPL), Barrow, Michael; Burns, Steven; and Kastner, Ryan, 2018. There are small changes in format and phrasing as a chapter within this larger paper. The dissertation author was the primary investigator and author of this paper.

Chapter 4

MR Elastography Data Driven Patient Specific Image Guidance

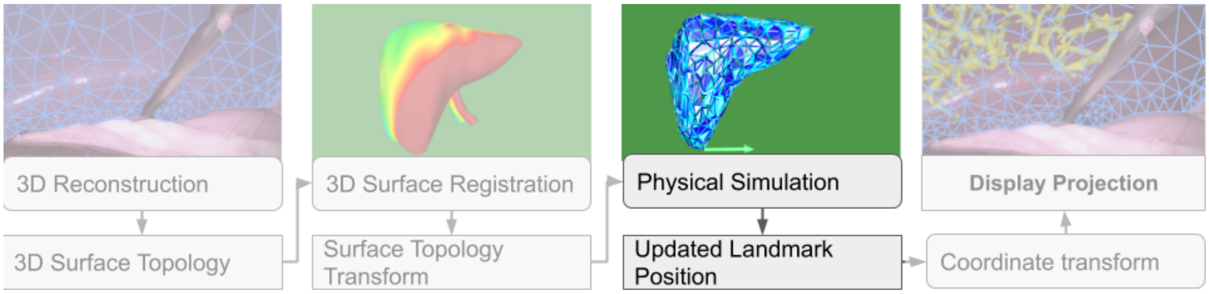


Figure 4.1: Our MRE physical simulation method improves efficiency of the online landmark placement step in a CAIG pipeline.

Recall the four major online steps required for accurate CAIG outlined in chapter 1. In this chapter we describe a "MR Elastography data driven method for physical simulation". This method applies to the physical simulation step. Figure 4.1 highlights the simulation step in the CAIG pipeline to illustrate where our contribution is useful within a larger CAIG system. Our data driven method improves the accuracy of landmark placement whilst maintaining the speed of the step. It therefore improves efficiency of CAIG.

4.1 Introduction

Preoperative MRI scans image the internal structure of solid organs and are essential for preoperative planning, image guidance during surgery and educating future surgeons. A drawback is that as a procedure progresses and the organ shape changes, the MRI landmarks may not remain relevant.

In this study we introduce an accurate patient specific physical simulation method for tracking surgical landmarks in liver procedures.

Figure 4.2 shows how MRI landmark positions can be interactively updated using physical simulation and how this can enhance surgical workflows.

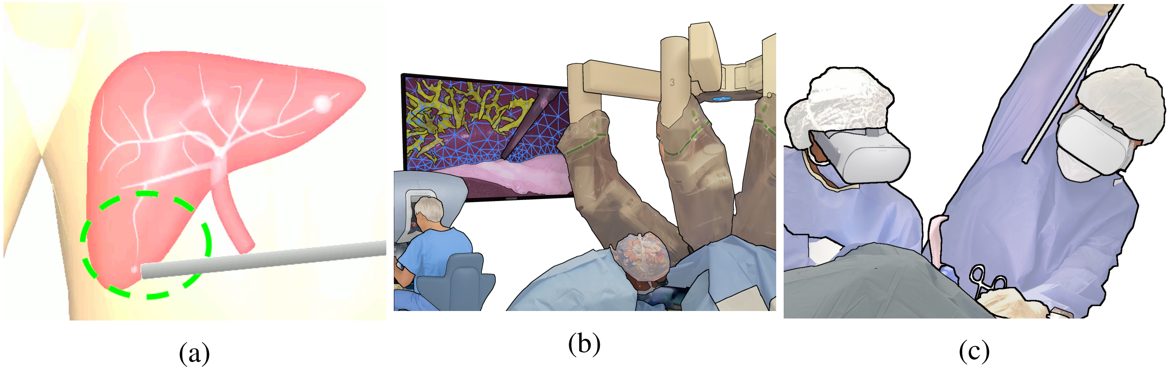


Figure 4.2: Illustration of physical simulation used in the surgical workflow. Figure 4.2a: In interactive planning tools, simulated tissue movements allow the surgery to be rehearsed to avoid complications during the actual procedure. Figure 4.2b: In image guidance applications, simulated tissue overlays help the surgeon track the position of surgical landmarks as the procedure progresses and the surgical scene changes from its appearance in preoperative MRI. Figure 4.2c: In surgeon training applications, simulated patient models can be used to practice surgical techniques many times over without risk to patients.

Figure 4.2a: During planning, simulations show how landmarks will be displaced by tools, allowing the surgeons to test various angles of approach for the best chance of success. Figure 4.2b: During procedures, simulation can track landmarks of interest, allowing the surgeon to avoid unnecessary damage to large vessels and keep track of tumor locations. Figure 4.2c: In an educational setting, simulation can allow inexperienced surgeons to practice procedures without risk to patient safety.

The current state of the art in liver simulation makes basic assumptions about the liver's mechanical properties. In particular, existing mechanical models assume a uniform stiffness or an *atlas* value. However, liver stiffness, a surrogate marker of liver fibrosis, tends to increase in the setting of chronic liver disease and may be heterogeneously distributed throughout the liver [181]. One diseased liver can be many times stiffer than another since patient to patient liver stiffness varies from 2 to 80 kPa [75]. Traditional simulations can result in inaccurate calculation of landmark displacement by surgical force since the atlas stiffness is not patient specific. Misplaced image guidance landmarks confuse surgical decision making and diminish the clinical usefulness of simulators. What is needed is a method to incorporate and model liver stiffness variation during surgical landmark placement. In order to accurately track landmarks and effectively aid surgical decision making, patient-specific liver stiffness and its heterogeneity must be considered.

Herein, **we describe the first surgical simulation able to model heterogeneous stiffness using biomechanics measured in vivo**.

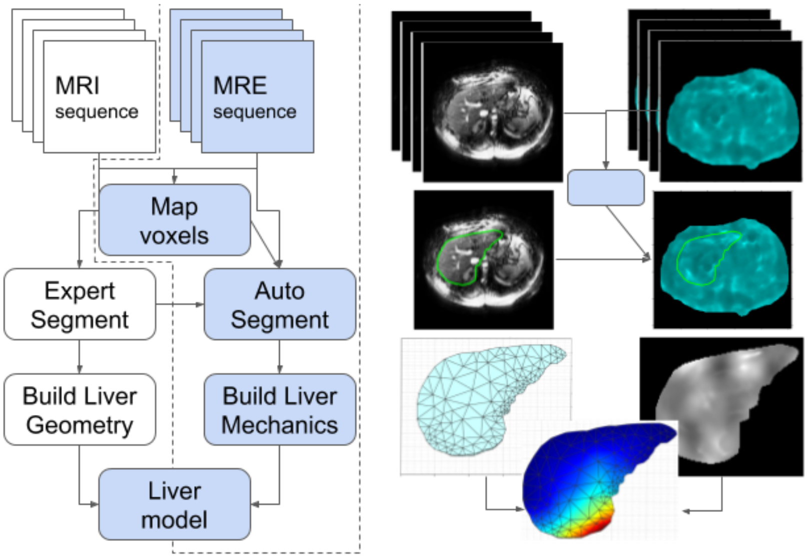


Figure 4.3: Proposed MRE biomechanical liver modeling method. Patient specific liver models are used by physical simulators to track surgical landmarks during surgical planning and image guidance. The MRE flow (highlighted in blue) has low workflow overhead since the MRI scans used in conventional modeling are acquired when generating MRE scans. A heterogeneous liver model is built from MRE and the natural MRI⟺MRE mapping is used to enrich a traditional MRI geometry model with our patient biomechanical model.

Our method uses liver stiffness maps measured using Magnetic Resonance Elastography (MRE) to position surgical landmarks. Figure 4.3 is a block diagram outlining our proposed liver modeling flow.

Figure 4.3 extends traditional liver modeling methods (white) with patient liver stiffness data (blue). After detailing our modeling method in section 4.2, we report surgical simulation studies of its clinical usefulness in sections 4.3 to 4.6.

Section 4.2: Patient specific MRE simulation method. We detail the contribution of our surgical simulation method and outline related work. Our approach is uniquely able to model patient liver stiffness, is computationally efficient and does not introduce significant overhead to the surgical workflow.

Section 4.3: Statistical cohort study of liver stiffness. We observe a great deal of stiffness variation in MRE scans, which indicates that MRE data is more representative of patient liver stiffness than an atlas.

Section 4.4: Patient simulation study. We simulate a simple liver procedure on our cohort to understand if MRE based landmark placement is significantly different to the traditional atlas method. We find clinically significant accuracy improvements using the MRE method.

Section 4.5: Image guidance study. We construct an image guidance platform and a silicone phantom to understand if MRE simulations are sufficiently accurate for image guidance. We find accuracy improvements over state of the art in a phantom palpation procedure.

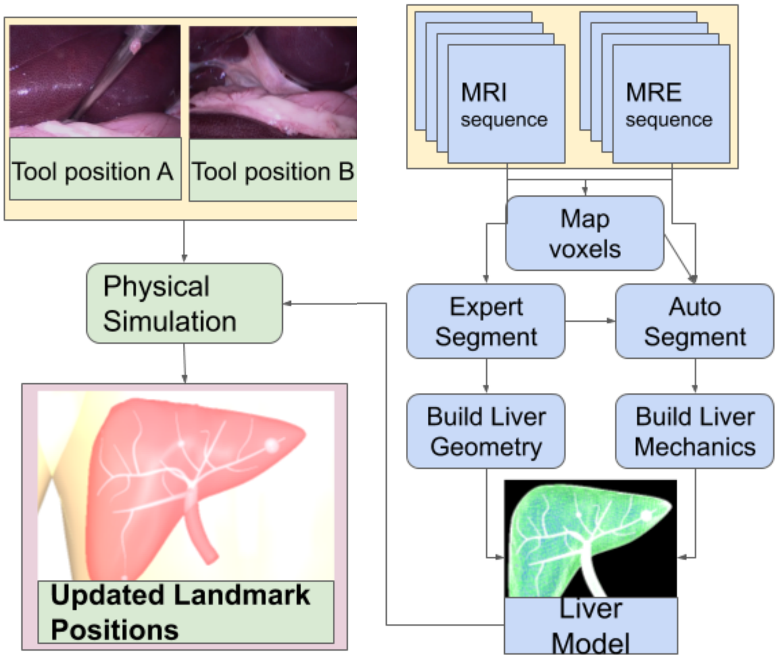
Section 4.6: MRE Validation study. We validate our MRE protocol accuracy by comparing our silicone phantom MRE stiffness with a traditional material characterization of the phantom. We find close agreement between the two sets of measurements.

The studies show that patient specific liver stiffness modeled using MRE based physical simulation is accurate, clinically significant, and practical in surgical landmark placement applications.

4.2 Patient Specific MRE Simulation method

Physical simulation is a two-step process. First, a liver model is created that defines the shape (geometry) and behavior (biomechanics) of the simulated liver. Secondly, landmark movement caused by surgical forces is calculated using numerical optimization of a system of equations governed by the behavioral model. In this section we first detail our research contribution to the modeling step. Next, we describe the simulation method we used to validate landmarks positioned with our patient specific models. Finally, we provide supplemental context with an overview of MRE and a discussion of related work that our method built on.

Figure 4.4 illustrates our patient specific implementation of the two-step physical simulation method.



(a) online steps (b) offline steps

Figure 4.4: Our patient specific simulation method. Input is yellow, output is red. Preoperative steps are blue, online steps are green. Patient MRE and MRI scans are fed into our patient specific liver modeling method to build a liver model in offline steps. Our model is integrated with an efficient meshfree simulator. A surgeon can input various tool forces and positions using a simulation framework that can be configured for use in either planning or image guidance workflows.

Our method is more accurate than the state of the art, does not introduce significant overhead to existing simulation workflows, and is compatible with the SOFA simulation framework, which has been used to enhance surgical decision making in liver procedures [182, 74, 136].

Offline Patient Specific Modeling

Our offline modeling tasks from Figure 4.4b are implemented as follows. First, the gross liver anatomy must be segmented from the MRI scans. In our case, this was done manually by an imaging expert, although semi-automatic and automatic methods are also possible [183, 184, 185, 52].

Next, these segmented images are turned into a 3D geometry during the “Build Liver Geometry” step. We use a fast meshing algorithm to define a watertight surface over the segmentation boundaries which are discontinuous due to scan slice separation. We used Deng’s techniques for triangulation since they allow good control over triangle size and have been shown to generate good quality meshes [186, 187]. We then compute a volumetric geometry using a 3D Voronoi decomposition of the surface mesh. Two geometric models are required for a liver model. The first is the volumetric geometry that specifies Degrees of Freedom (DOF) in our particle based meshfree simulation. The second type of geometry is the “contact geometry” which is implemented as a mesh. A contact geometry is used to specify the outer extremities of the liver and provides a mapping of interactive forces to the mechanical model. The contact geometry is required since the Voronoi decomposition does not neatly follow the liver geometry edge.

Once the geometry has been defined we build the patient specific biomechanical model. In the “Auto Segment” step, each DOF centroid in the Voronoi decomposition of the geometry is assigned an MRE voxel using the relationships as defined in the DICOM metadata (“Map Voxels”). We match the centroid to the MRE data by minimizing L2 distance to find the nearest MRE voxel neighbor to each Voronoi centroid. Next, in the “Build Liver Mechanics” step, the

MRE shear modulus is converted to a Young's modulus using Equation (4.1):

$$E = G(2(1+v)) \tag{4.1}$$

where G is the shear modulus, E is the Young's modulus, and v is the Poisson's ratio.

The patient specific liver models are then converted to a SOFA specific node data format. Our liver model can then be used in planning, image guidance, or training surgical simulations.

Online Simulation

Our patient specific method uses meshfree physical simulation to implement Figure 4.4a "Physical Simulation". Meshfree approaches are well-suited to surgical simulation because they do not require re-meshing after topology changes (e.g., cutting) that occur during a resection and have been previously implemented in SOFA [188]. Our MRE simulation uses the following canonical semi-discrete formulation:

$$\mathbf{M}\ddot{\mathbf{q}} - \mathbf{f_{int}}(\mathbf{q}, \dot{\mathbf{q}}) = \mathbf{f_{ext}}(\mathbf{q}, \dot{\mathbf{q}}) \tag{4.2}$$

where $\mathbf{q}$ is the nodal displacement vector, $\dot{\mathbf{q}}$ is the nodal velocity vector, $\ddot{\mathbf{q}}$ is the nodal acceleration vector, $\mathbf{M}$ is the mass matrix, $\mathbf{f_{int}}$ is the internal force vector, and $\mathbf{f_{ext}}$ is the external force vector.

We choose a computationally efficient discretization of this system using the *implicit Euler* method [189] which creates a system of linear equations. The linear approximations allow us to achieve video rate real time performance in our simulations. The system of equations is fed into the conjugate gradient descent method that will ultimately solve for $\delta\dot{\mathbf{q}}$ for each time step. The implicit Euler method is adopted as follows:

$$(\mathbf{M} - h\mathbf{C} - h^2\mathbf{K})\delta\dot{\mathbf{q}} = h(\mathbf{f_{ext}} + h\mathbf{K}\dot{\mathbf{q}}) \tag{4.3}$$

where h is the time step (typically in milliseconds for our studies), $\mathbf{K}$ is the Jacobian of first order derivatives of $\mathbf{f_{int}}$ with respect to $\dot{\mathbf{q}}$ (also known as the stiffness matrix), and $\mathbf{C}$ is the Jacobian of first order derivatives of $\mathbf{f_{int}}$ with respect to $\mathbf{q}$ (also known as the damping matrix). Note that K, C, and M may be precomputed offline and that this corresponds to our "Offline Patient Specific Modeling" step using MRI and MRE inputs in Figure 4.4b.

The physical simulation step in Figure 4.4a is computed numerically via the following mappings:

$$A = (\mathbf{M} - h\mathbf{C} - h^2\mathbf{K}) \tag{4.4}$$

$$\mathbf{b} = h(\mathbf{f_{ext}} + h\mathbf{K}\dot{\mathbf{q}}) \tag{4.5}$$

$$\mathbf{x} = \delta\dot{\mathbf{q}} \tag{4.6}$$

such that $A\mathbf{x} = \mathbf{b}$, where $\mathbf{x}$ contains the vector encoding of surgical landmark positions to resolve for the current time step. Once $\delta\dot{\mathbf{q}}$ is solved for each node in the system, calculating the change in global position, $\delta\mathbf{q}$, is simply done by integration. Computing $\delta\dot{\mathbf{q}}$ is the computational bottleneck. First there is an initial guess: $\mathbf{x_0}$

$$\mathbf{r_n} = \mathbf{b} - A\mathbf{x_n} \tag{4.7}$$

$$\alpha_n = \frac{\mathbf{r_n^T}\mathbf{r_n}}{\mathbf{p_n^T}A\mathbf{p_n}} \tag{4.8}$$

$$\mathbf{x_{n+1}} = \mathbf{x_n} + \alpha_n\mathbf{p_n} \tag{4.9}$$

$$0 \leq n \leq N, n \in \mathbf{Z} \tag{4.10}$$

(When $n = 0$, $\mathbf{p_0}$ is initially set to $\mathbf{r_0}$)

The steps above illustrate a portion of one iteration of the conjugate gradient descent method with each subsequent $\mathbf{x_n}$ representing a closer approximation of $\mathbf{x}$. The linear solver continues to iterate until the loop is terminated upon reaching preset error bounds. In our studies

$n = N = 200$. Since MRE data is used to solve Equation (4.4) once offline, our patient specific simulation method is computationally efficient and does not exacerbate the conjugate gradient bottleneck.

Magnetic Resonance Elastography (MRE)

Our simulation method uses patient specific MRE data to model stiffness. MRE measures mechanical properties of the liver as *heterogeneous voxel stiffness maps*. These maps allow our simulator to more accurately model the liver's response to forces as we show in our experimental results.

MRE is a non-invasive magnetic resonance method of measuring tissue mechanical properties in vivo [190]. Its primary clinical use in the liver is as a surrogate noninvasive measure of liver fibrosis, which is important in diagnosis and evaluation of chronic liver disease in cirrhosis [191]. It is also an emerging tool for monitoring tumors and changes in other soft tissues such as the heart, brain, and breast [192].

MRE induces a shear wave into the tissue using a mechanical driver. An acoustic driver is coupled to a non-ferromagnetic diaphragm placed on a patient inside an MRI machine. A series of gradient sensitized MRI images are taken while the driver maintains a shear wave in the patients tissue. A 2D or 3D elastic modulus map is derived by performing direct wave inversion, local frequency estimation, algebraic inversion of the differential equation, or data-driven hybrid techniques including numerical optimization and machine learning approaches [193, 194, 195].

MRE has been shown to be an accurate noninvasive tool for measuring stiffness [196, 197, 198, 199, 200, 201]. Thus, MRE is a promising candidate to enhance simulation accuracy; however, further studies are needed to understand the clinical significance of MRE in surgical planning and image guidance applications. We perform those studies in Sections 4.3 to 4.6.

Surgical Simulation

The liver poses many challenges to real time physical simulation. It is an organ of soft, hyperelastic, and heterogeneous materials that is difficult to model in real time. Various approaches to the problem have been proposed and our method builds on this body of work.

Early mass spring models sought to maximize computation efficiency, but as hardware has become more powerful, a myriad of specialized CPU and GPU Finite Element Analysis (FEA) solvers have been implemented [202, 203, 204] and more complex models have been proposed. Zhang's survey provides a broad overview of surgically oriented models that reviews both common and exotic models for real time simulation in surgery [118]. Most research efforts have been focused on advanced FEA methods that mitigate the issues of large displacements that occur in tissue during surgery [205, 206, 207]. Researchers have also explored approaches for simulating the topological change caused by surgical tools. A good summary of cutting simulation methods is Wu's recent survey [208] which provides a taxonomy of advances in computationally efficient FEA and non FEA methods not covered by Zhang's survey. Because of the complexity of surgical scenes, models and simulations are typically packaged in frameworks of simulation libraries which are then configured to perform surgical planning, image guidance, or training simulation. Frameworks include "SPRING" [209] framework which uses a mass spring model, SOFA [210] which supports FEA and meshfree simulation, and Nvidia FleX which uses position based dynamics [211]. Of these frameworks, SOFA is the most mature and feature rich.

With the improvement in simulator capability, researchers have invested in creating patient specific simulations for accurate image guidance. An advanced approach is described in Haouchine's work which considers heterogeneity of tissue mechanics and anisotropic materials within the liver [69]. Hauochine's work, however, focuses mainly on geometry and assumes stiffness coefficients from an atlas. Plantefeve's also performs patient specific liver modeling and uses offline segmentation approaches to create separate models of liver blood vessels and parenchyma [70]. Plantefeve mitigates the unrepresentative nature of atlas stiffness data by attempting to modify the elastic modulus of the model to reach convergence with the simulation and video feed online. However, this approach does not take into account heterogeneity of the parenchymal stiffness and modifying the modulus may yield only superficially better outcomes as a result.

Our patient specific method develops the state of the art with the first model of heterogeneous liver stiffness built using in vivo measurements. In the remainder of the paper we compare our work with the prior art to understand the clinical significance of our method.

4.3 Liver Stiffness Variability Study

To help quantify the benefits of using patient specific MRE data for physical modeling, we perform a series of experiments to determine if and how our model will affect the results of the physical simulation. This section describes the first of these experiments, which performs an analysis on a set of patient MRE scans. Here we seek to understand if patient liver stiffness varies significantly from a traditional atlas value, thus validating the need for patient specific modeling. Although liver stiffness variation caused by disease has been reported between 2 – 80KPa [75], it is not clear how the distribution of cirrhotic tissue impacts surgical image guidance landmark placement. As we describe in this section, our patient group shows large inter-organ stiffness variations, which motivates the need for patient specific modeling. We describe these results in more detail in the following.

MRE Stiffness Analysis Methods

We used two studies to analyze liver stiffness variation. First, we studied inter-patient stiffness variation. Our goal was to understand if the gross liver stiffness of our patient cohort was represented well by an atlas stiffness. If not, this indicates that it could be valuable to model patient specific liver stiffness in surgical image guidance. Second, we studied intra-organ liver stiffness variation. We aim to understand if stiffness variations within the liver were well represented by a single stiffness value for the entire liver. If not, modeling patient liver stiffness with MRE is interesting as elastography is the only known method of measuring heterogeneous stiffness maps in vivo. Our statistical experiments were integrated into the proposed flow of our MRE physical modeling method as illustrated in Figure 4.5.

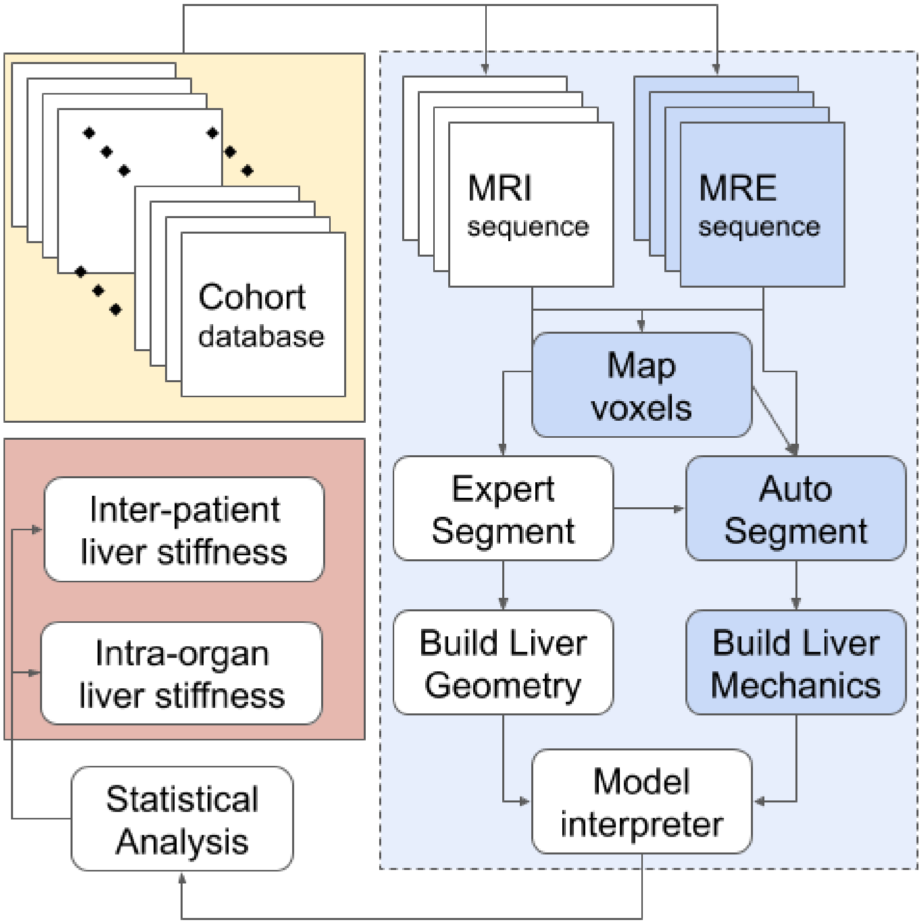


Figure 4.5: Cohort liver stiffness statistical study experiment. The input to our analysis is a cohort database of MRI and MRE scans (highlighted by the yellow box). Our method builds patient specific biomechanical models using this MRE data (blue box). A statistical analysis is performed on models of the entire cohort. Our inter-patient and intra-liver statistics are outputs indicated by the red box.

For each of our studies we repeat similar steps for every patient in our cohort database. First, our liver geometry models are built using expert segmentations from liver MRI to extract the surface of gross liver anatomy. Next, we use the mapping between MRI and MRE voxels to segment liver stiffness voxels from MRE. A patient specific biomechanical model can be built from the geometry and stiffness data at this point.

The next steps differ for each of our stiffness studies. In our inter-patient analysis, we use a model interpreter to compute the mean stiffness ($\bar{E}$) of a patients liver. Once we have $\bar{E}$ for the entire set of cohort scans, we calculate the global mean and probability distribution which can be compared directly with the atlas stiffness cohort study [212].

In our inter-organ analysis, we use a model interpreter to compute the standard deviation (σ) of the patients liver stiffness. Once we have σ for the entire cohort, we cluster them around histogram bins for interpretation. For example, setting bin width to the $p = 0.95$ of the atlas stiffness study distribution allows us to analyze how representative a single liver stiffness value is likely to be.

Results and Discussion

We performed a cohort study using 119 2D MRE scans acquired between March 2012 and December 2013 under a Health Insurance Portability and Accountability Act (HIPAA) compliant study. Our cohort sample population was adults (>18 years of age) who were severely obese (body mass index >35 kg/m^2) and were being evaluated for weight loss surgery. The exclusion criteria were contraindications to MRI or history of known liver disease besides potential nonalcoholic fatty liver disease. MRE were acquired using a GE Medical Systems Discovery$^{\text{TM}}$MR750 system at 3.0T field strength using a spin echo pulse sequence and the standard abdominal MRE protocol. Voxel resolution was 1.64×1.64mm, slice thickness was 10mm, and slice spacing was 10mm. A weight-based does of gadobenate dimeglumine at a standard bolus rate of 2 ml/second was administered to patients by peripheral intravenous catheter.

Inter-Patient Liver Stiffness Variability Study

Figure 4.6 groups the mean stiffness of each patient's liver ($\bar{E}$) into a histogram.

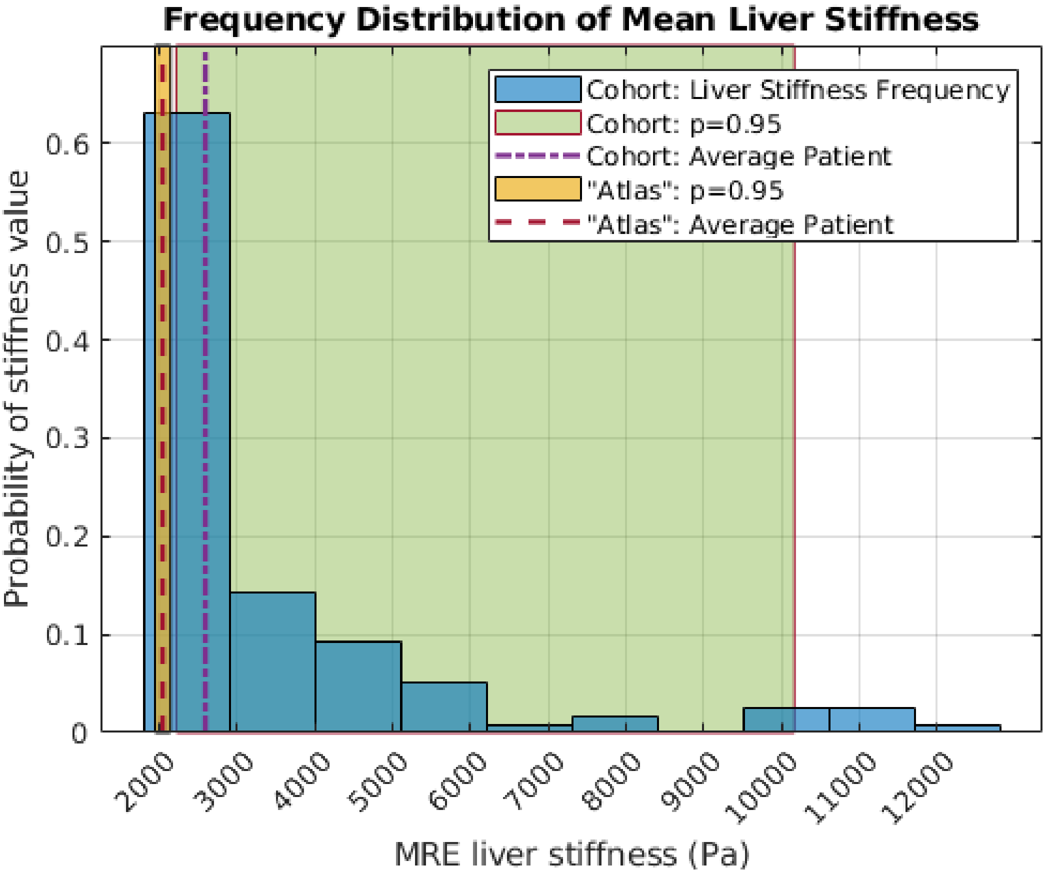


Figure 4.6: Frequency distribution of $\bar{E}$ liver stiffness in our cohort as a histogram. The average patient of our cohort had $\bar{E}$ within 0.5kPa of the average patient $\bar{E}$ from the atlas cohort. However, our liver stiffness distribution, $p = 0.05$ (green), varied a great deal from the atlas $p = 0.95$ (yellow).

Comparing our cohort statistics with the atlas cohort from [212] reveals a long tail in $\bar{E}$ frequency that can't be represented by atlas data. Although we saw close agreement (<0.5kPa) between the average patient $\bar{E}$ in our cohort and in the atlas cohort (2.1 and 2.6kPa, respectively), the $\bar{E}$ frequency distribution of our cohort varies largely from the atlas cohort. This variation can be seen by considering the green and yellow regions in fig. 4.6 which represent $\bar{E}$ for $p = 0.95$ (95%) of our cohort and the atlas cohort respectively. These regions do not overlap and it is unclear that the two distributions are related enough to reliably use atlas data in simulation. For example, $\bar{E}$ at the upper limit of $p = 0.95$ in our cohort is 4.8 times stiffer than $\bar{E}$ at the upper limit of $p = 0.95$ in the atlas cohort.

Intra-Organ Liver Stiffness Variability Study

Figure 4.7 graphs liver heterogeneity or; intra-liver stiffness standard deviation (σ). σ is overlaid on our cohort $\bar{E}$ frequency distribution and illustrates mean σ for each $\bar{E}$ bin of fig. 4.6.

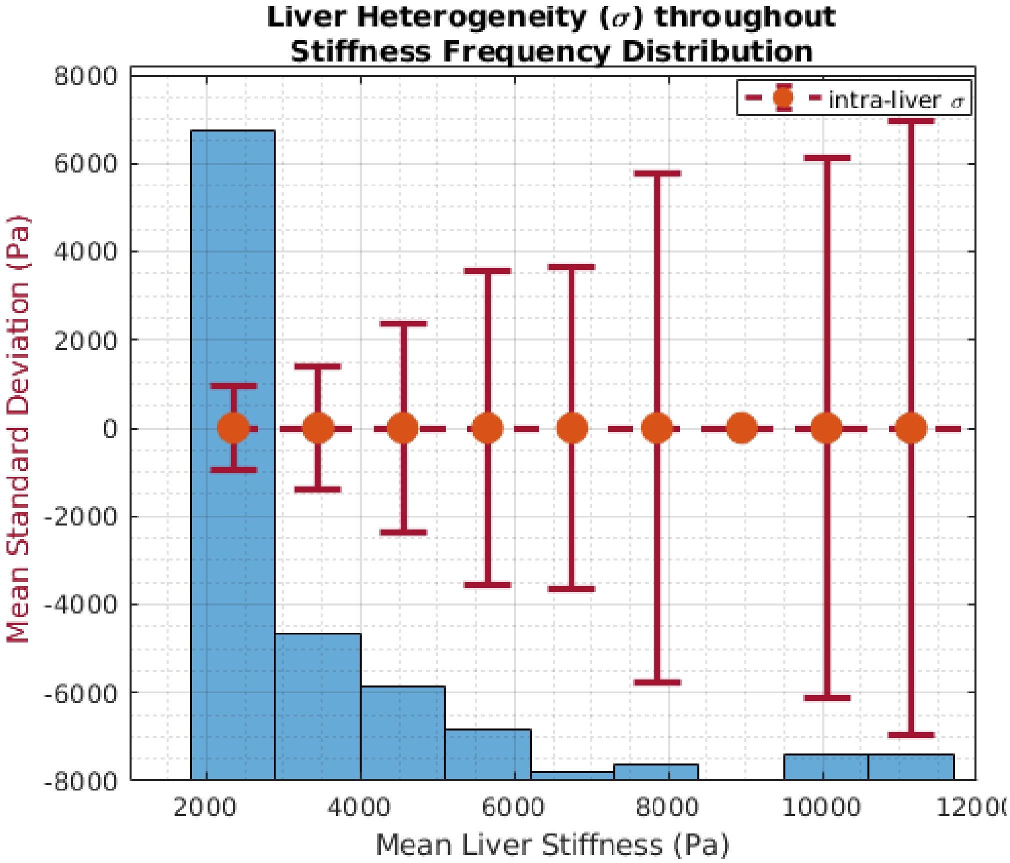


Figure 4.7: Intra-organ stiffness variation. As livers become more stiff or diseased, heterogeneity increases.

The first data point from fig. 4.7 is where $\bar{E}$ in our cohort agrees closely (±0.5kPa) with $\bar{E}$ of the atlas MRE data set and represents the majority (64%) of our cohort. For these cases, intra-organ σs is larger than the atlas $p = 0.95$ interval from fig. 4.6 and is 0.95kPa vs 0.5kPa.

This indicates a lot of valuable heterogeneous stiffness information must be discarded to create the single atlas stiffness value used in traditional simulations even in cases where the atlas value corresponds well to average stiffness of a patient's liver. Since σ increases monotonically with $\bar{E}$, the importance of using heterogeneous MRE stiffness data increases with disease progression.

In summary, variation in inter-patient and intra-organ stiffness highlights that atlas data does not accurately represent a large fraction of our cohort and motivated our study on the clinical significance of MRE based physical simulation.

4.4 Clinical Significance Study

In section 4.3 we found significant stiffness variation in a cohort of patient liver MRE scans. Since stiffness determines how simulators position image guidance landmarks, it is important to understand the clinical implication of patient specific biomechanics in image guidance applications. We designed two simulation experiments using the SOFA [213] surgical simulation framework to study the effect of our patient specific simulation on image guidance accuracy.

Simulation Analysis Methods

We study how patient specific liver stiffness impacts vessel avoidance, where a surgeon seeks to avoid damaging major vessels. Figures 4.8 and 4.9 are block diagrams of our two studies. In general we perform our experiments by applying surgical forces to a physical simulation of a liver model. Yellow boxes are input, blue are offline preoperative steps, green are online steps and red boxes are output.

Inter-Patient MRE Liver Stiffness Study

Figure 4.8 illustrates how we emulated liver retraction, a common surgical procedure to expose the rear of the liver, with a virtual Nathanson retractor load applied to segment VI of the

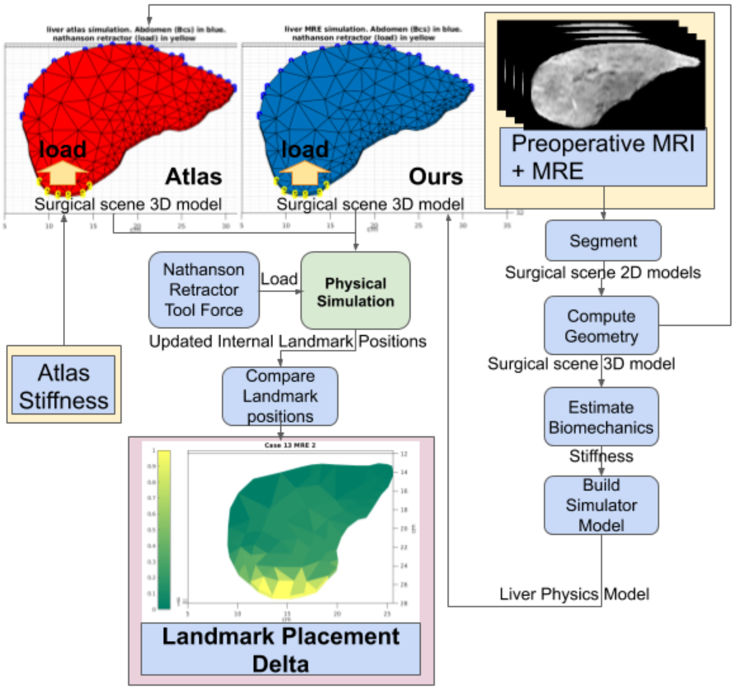


Figure 4.8: Inter-patient stiffness variability convergence experiment. First, liver geometry was segmented. Next, a patient specific liver physical model was built using mean MRE stiffness. Standard boundary and load conditions were applied to the models and a retraction load was applied. Finally, L2 distance between the simulations was computed to determine clinical significance of liver stiffness heterogeneity.

liver. We applied the load to liver physics models built using our MRE method and compared those simulations to a traditional atlas based method [188]. Our comparison measured the L2 distance between surgical landmark placements in both simulations. Displacement delta over 5mm was considered clinically significant [44].

Intra-Organ MRE Liver Stiffness Study

Figure 4.9 depicts our intra-organ stiffness study which compares landmark placement differences caused by stiffness variation within the liver. Common simulation parameters are given in table 4.1.

We regularize geometry, boundary and load conditions such that liver stiffness is the only variable in the convergence study. The steps of our liver heterogeneity study are:

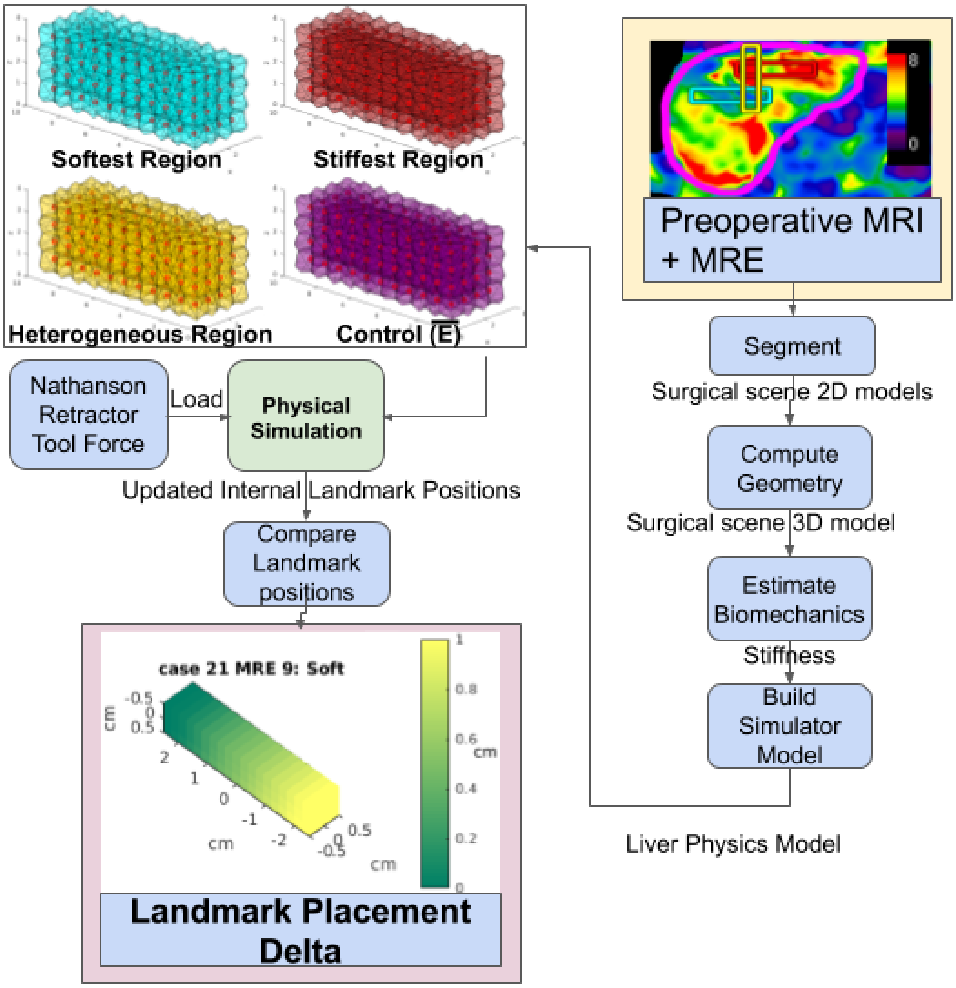


Figure 4.9: Intra-organ stiffness variability convergence experiment. First $\bar{E}$ for the patient is computed as a control. Next, MRE scans are searched for stiff, soft and the most heterogeneous stiffness regions. Next, standard boundary and load conditions are applied. Finally L2 distance between the regions and a control beam was computed to determine clinical significance of liver stiffness heterogeneity.

Table 4.1: Region of interest simulation parameters used in stiffness variability case study. Boundary conditions model stiffness of the abdominal wall [7].

Length	Width	Height	ρ	Boundary Conditions	Retraction Force
50mm	10mm	10mm	997kg/m3	22kPa	0.05N

First, we create a control beam using the entire liver MRE as shown outlined by magenta in the input box of fig. 4.9. The control beam has a uniform stiffness of $\bar{E}$, or the mean stiffness of the whole liver. Next, we locate the stiffest, softest and most heterogeneous regions of stiffness within a patient MRE as indicated by red, yellow and cyan rectangles marked on the MRE input of fig. 4.9.

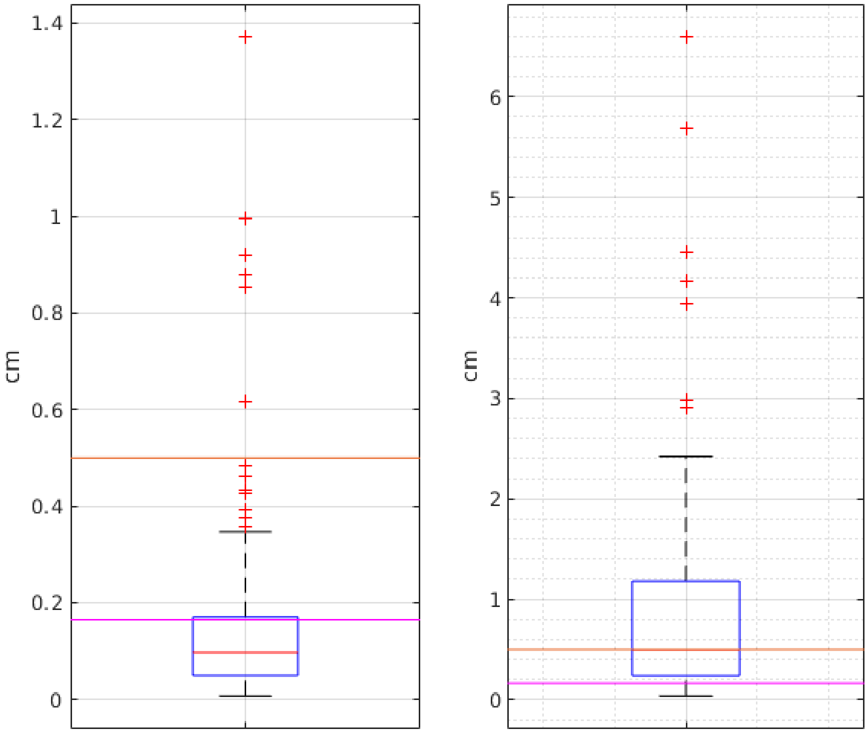
We then construct a meshfree beam simulation for the four regions using the SOFA surgical simulator framework [210]. A hoisting "retraction" load is applied to each region. This simulates a common maneuver in liver surgery to expose the rear of the liver to the surgeon. Finally, we perform a registration of the surface "contact" geometry of the three regions to the control. L2 Strain difference of 5mm is considered clinically significant disagreement with $\bar{E}$ [44] and if exceeded would confirm that liver heterogeneity must be simulated for accurate guidance landmark placement.

Results and Discussion

Inter-Patient MRE Liver Stiffness Study

We ran the liver retraction experiment illustrated by fig. 4.8 on 102 of our cohort scans. Figures 4.10 and 4.12 are a summary of our findings.

Figure 4.10 shows the trends of L2 landmark placement delta for our Inter-patient liver stiffness study (fig. 4.8). First the average mean displacement difference in fig. 4.10a is close to the MRE voxel resolution of 1.65×1.65 (marked as a magenta line on the box plots). The close agreement suggests that using a stiffness atlas has merit. However the average displacement difference around the virtual retractor in fig. 4.10b has poorer agreement. The $5mm$ mean difference is clinically significant for image guidance.

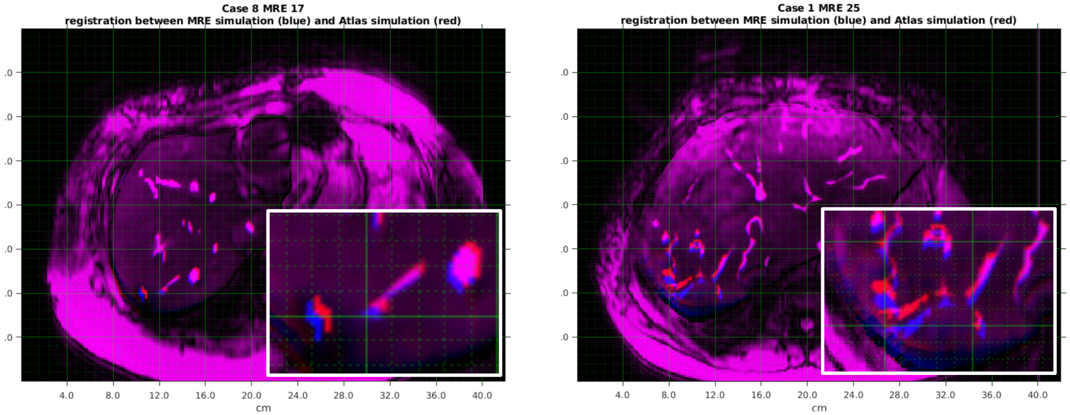


(a) mean atlas error (b) max atlas error

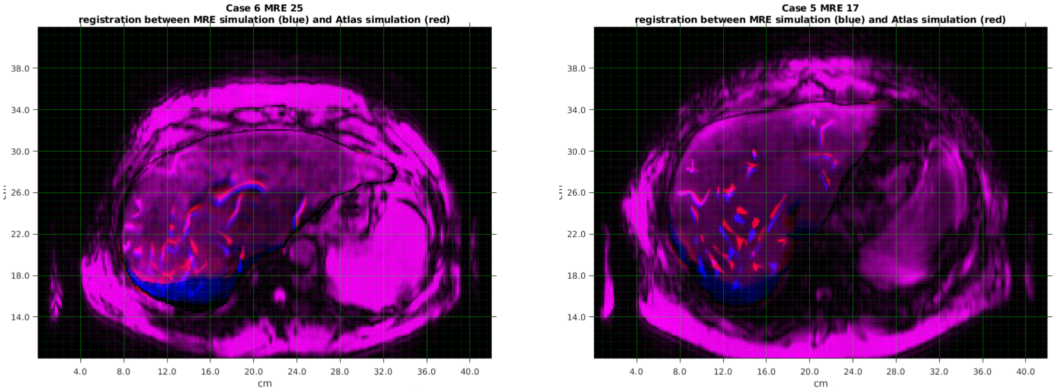
Figure 4.10: Inter-patient landmark placement error statistics for liver retraction. Orange is clinically significant and magenta is imperceptible error. Mean squared error (fig. 4.10a) is misleadingly low. Max error (fig. 4.10b) is often significant and occurs at the retraction site where accuracy is critical.

Figure 4.12 illustrates the image guidance implications of the cohort L2 statistics from fig. 4.10.

We show four cases that represent different statistical regions of L2 landmark placement error from the box plot of fig. 4.10b. Each image is a composite of an MRE data-driven simulation (in blue) and an atlas based simulation (in red). Where the simulations agree, the image appears purple. Where they disagree, the image appears more blue or red. The image guidance application is vessel location and some larger vessels in the liver have been CT enhanced as image guidance landmarks. The high contrast of the highlighted vessels enhances blue or red color shifts. Vessel placement is noticeably different in all cases and ranges from ≈ 5mm to ≈ 10mm. The disagreement is clinically significant and motivates the use of MRE in image guidance.



(a) Mean strain δ with error zoom (b) +1 standard dev strain δ with error zoom



(c) +2 standard dev strain δ (d) strain δ outlier

Figure 4.12: Simulations showing the effects of inaccurate atlas stiffness on vessel avoidance. MRE and atlas retraction simulations are blue and red channels in each image respectively. Landmark misplacement is indicated by differences in image color. Figure 4.11a is a typical case and error is clinically significant. 50% of our simulations found worse. Figures 4.11b to 4.11d are example cases from each standard deviation above the mean in fig. 4.10b and show how far error can spread throughout the liver.

Intra-Organ MRE Liver Stiffness Study

We ran the heterogeneity study illustrated by fig. 4.9 on 119 of our cohort scans. We extracted the stiffest, softest and most heterogeneous stiffness regions from MRE to study the variability of surgical landmark placement in image guidance caused by liver heterogeneity. A retraction load was applied to the three regions and a convergence study was used to compare simulated landmark placement with a homogeneous $\bar{E}$ control. Figure 4.14 is a sample of experimental results while fig. 4.13 graphs the set of cohort intra-liver landmark placement variation.

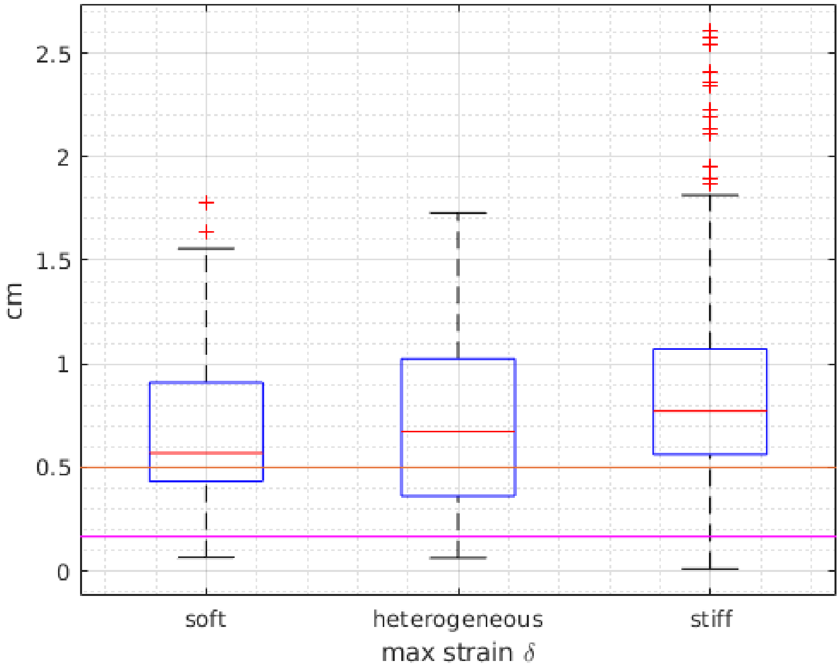
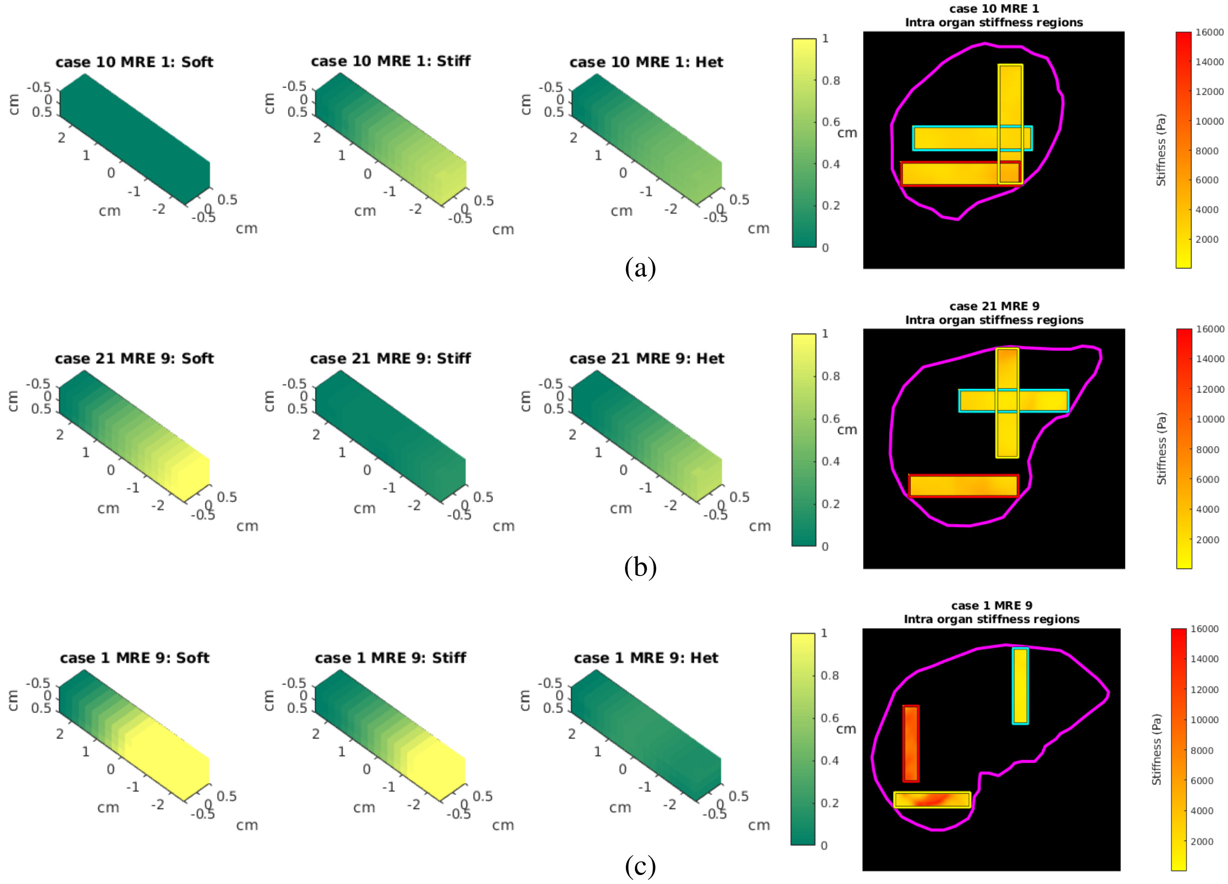


Figure 4.13: Intra-liver landmark placement error statistics. Orange is clinically significant and magenta is imperceptible. Every internal region has clinically significant deviation from the $\bar{E}$ control most of the time. This figure shows that simulating a single liver stiffness limits accuracy and heterogeneity must be considered.

Figure 4.13 is a box plot of L2 displacement difference between a control and heterogeneous simulation of liver retraction. In the control, a patient specific but homogeneous MRE liver stiffness of $\bar{E}$ was used. In the heterogeneous simulations, either the soft, heterogeneous, or stiff regions of the liver were simulated. The magenta line is minimum perceivable L2 difference of one voxel and the orange line is clinically significant landmark placement difference. Each stiffness region has its red line above the orange line, indicating clinically significant difference with the $\bar{E}$ control in over 50% of the study group. In other words, the intra-organ stiffness variation in our cohort made accurate modeling of the liver using one stiffness value impossible.

Figure 4.14 illustrates that even if the soft or heterogeneous or stiff region agree with an $\bar{E}$ simulation, variation in liver stiffness means accurate simulation with one stiffness value is still impractical. Figures 4.14a to 4.14c show cases where agreement with the $\bar{E}$ in one area was coincident with a large disagreement in all other areas. In fact, only 5% of our cohort showed clinically acceptable vessel placement agreement between a $\bar{E}$ simulation and one that accounts for intra-organ heterogeneity.

In summary, the orange line in fig. 4.13 shows that intra-organ heterogeneity caused



(a)

(b)

(c)

Figure 4.14: Sample cases from our Intra-Organ MRE liver stiffness study. We found accounting for heterogeneity was clinically significant to surgical landmark placement accuracy in 95% of our cohort. Figures 4.14a to 4.14c illustrate the problem of attempting to choose a single stiffness value for liver models, even when that value is patient specific. Simulation L2 distance (agreement) is depicted as a heat map from agreement in green to 1cm difference in yellow. We simulate liver retraction (hoisting) and compare landmark placement of the stiffest, softest and most heterogeneous liver regions to a mean ($\bar{E}$) control simulation. In each graphed case, we see landmark placement agreement between the control and one other extreme stiffness region of the liver, but clinically significant disagreement in all others. Our method is uniquely able to account for this heterogeneity.

clinically significant differences in landmark placement for most cases in our cohort convergence study and motivated our phantom validation study of MRE enriched image guidance.

4.5 MRE Image Guidance Validation Study

In our previous study (section 4.4) we found the difference in accuracy between our MRE method and the traditional atlas method was clinically significant. This is interesting for preoperative applications, however, a larger image guidance pipeline is needed to validate the accuracy of MRE enriched physical simulations in image guidance applications.

Simulation Validation Methods

We built a specially instrumented image guidance pipeline that implemented our physical simulation method on silicone phantom palpation procedures. Figure 4.15 shows our pipeline with inputs in yellow, output in red, online steps in green and offline steps in blue.

The general steps of our image guidance validation study using the apparatus in fig. 4.15 were: First, we gathered ground truth of phantom palpations using our instrumented pipeline. Secondly, we performed a physical simulation for use as image guidance on the canned palpation ground truth. Thirdly we measured the L2 distance error between our simulated scene and the ground truth data. Below we describe each of the three steps in more detail.

Gathering Ground truth

Our silicone phantom was placed beneath the palpation tool as depicted in the red box of fig. 4.15 This mimics the design of silicone validation experiments in the literature [74] and has an inline load cell to measure palpation force. Force data from load cell was used as force input to the simulation box in fig. 4.15. Next we took our first ground truth scan using the laser scanner. This was used to perform the initial registration in fig. 4.15 and find the initial coordinate

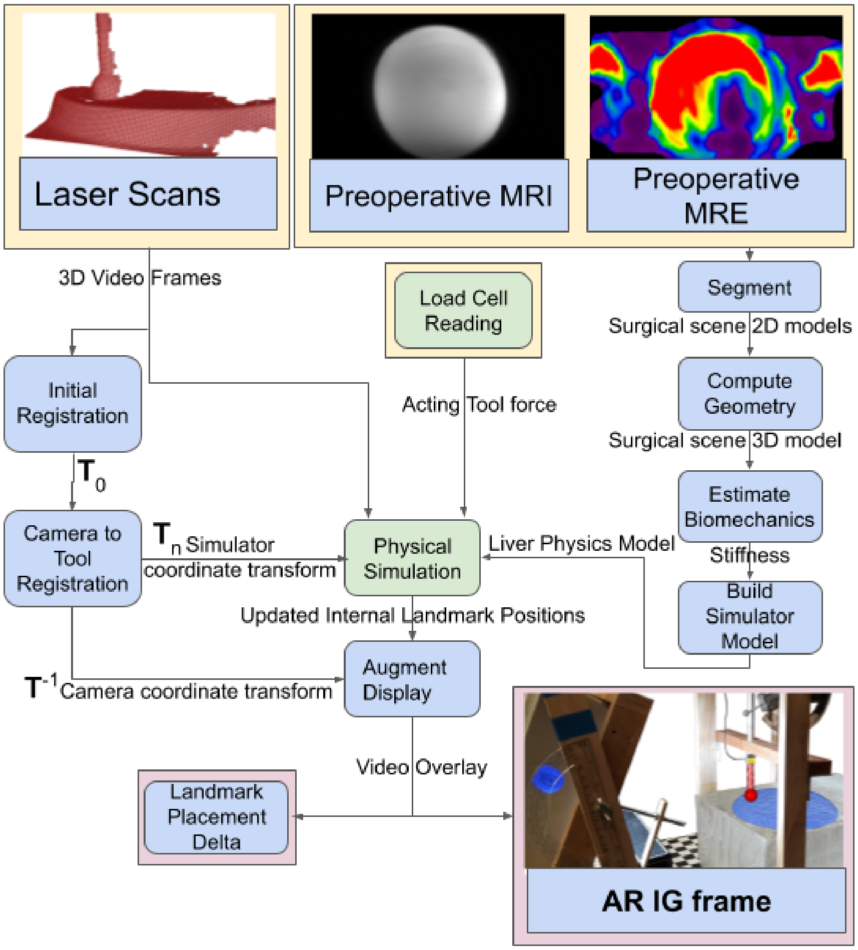


Figure 4.15: Phantom Palpation apparatus, our equipment for applying a load to a cylindrical phantom. A futek micro load cell records the load applied to a phantom via the round palpation tool using the USB data logger attached to the frame. A NextEngine laser scanner is positioned to record phantom indentation caused by the tool.

space transform between real world and simulation space: $\mathbf{T}_0 = [\mathbf{S}\ \mathbf{R_0}\ \mathbf{t_0}]$. We then applied a palpation load, recording force applied by the tool and taking ground truth scans of the phantom scene at as the load was ramped using a worm gear up to the limit of our load cell.

Performing physical simulation

We ran our MRE flow illustrated in fig. 4.3 and described in section 4.3 to build a "patient specific" model of the phantom using the following steps: First, we used segmented

MRI measurements to build a geometric model of the phantom. Dimensions of our phantoms are reported in table 4.2. Second, we used MRE and MRI scans of the phantom to build a biomechanical model. We then constructed the physical simulation depicted in the simulation box of fig. 4.15. We simulated the palpation using the SOFA framework and the same meshfree method used in our clinical significance study (section 4.4). Inputs to the simulation were the MRE phantom model built using our patient specific modeling method, ground truth force data and the coordinate transform ($\mathbf{T}_n$). Our simulator outputs volumetric tool and phantom position data which can be used to position surgical landmarks. We transformed our simulation to the ground truth laser scanner coordinate system using $\mathbf{T}_n^{-1}$ so it can be used to create a image guidance overlay. The concept is illustrated in the red box of fig. 4.15 by the blue phantom and red palpation tool image guidance overlays.

Validating MRE simulation accuracy

We first ran our simulator on palpation loads that corresponded with ground truth phantom surface scans. Next, we extracted the simulated surface of the surgical scene from output of our simulator. Finally, we calculated L2 distance between the ground truth scan and our simulation to evaluate the accuracy of an MRE image guidance pipeline.

Results And Discussion

We studied image guidance applications of our MRE method using “soft” (A), “stiff” (B) and “heterogeneous” (A/B) silicone phantoms. Each phantom was cylindrical and mounted on a torso “chassis”. Phantom geometry details are given in table 4.2. We used two medical grade silicone compounds to build our phantoms while the chassis was composed of ballistic gel. Details are provided in table 4.3.

We constructed SOFA models of our phantoms using our patient specific method. Figure 4.16 illustrates how our liver physics models and silicone phantoms were used in our image

Table 4.2: Dimensions of the silicone phantoms used to validate MRE enriched image guidance. The heterogeneous phantom (A/B) was bifurcated with two strata of different stiffness silicone.

Cylindrical Phantom		**Oblong Phantom Chassis**		
Diameter	Height	Length	Width	Height
225mm	26mm	350mm	300mm	150mm

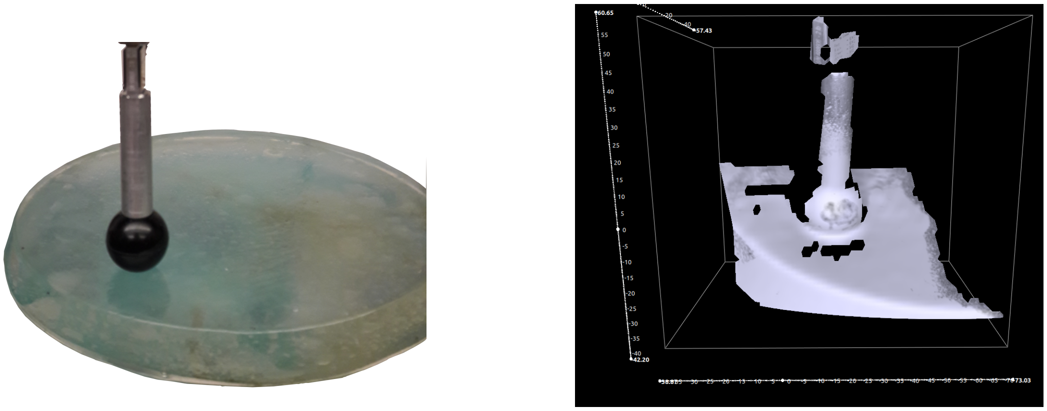
Table 4.3: Material composition of the silicone phantoms used to validate MRE enriched image guidance. "A" and "B" refer to the two gels with different stiffness used to create the homogeneous and heterogeneous phantoms.

	Compound A	**Compound B**	**Torso**
Manufacturer	Humimic Medical	Humimic Medical	Clear Ballistics
Product code	608729261421	852844007437	240A OR(10%)

guidance pipeline to validate our simulation method.

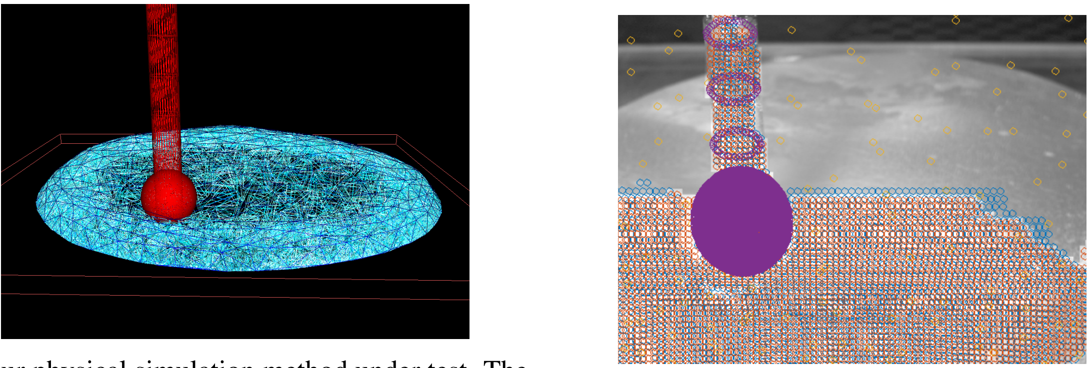
Figure 4.16a shows a silicone phantom during a palpation experiment. Palpation force was applied using the spherical tool tip interfaced with a micro "S" beam load cell which is the rectangle at the top of the image. First, the load cell was calibrated using an accurate desktop balance. Next, the palpation tool was abutted to the load cell and adjusted to be perpendicular to a phantom. The palpation force was then applied using a worm gear to depress the palpation tool into phantoms. Palpation force was recorded from the load cell throughout the palpation and used as the simulator force input ("Load Cell" in fig. 4.15). We performed a manual initial registration of the real world (scanner) and model (SOFA) coordinate systems, followed by ICP refinement to compute the $\mathbf{T}_0$ transform between the ground truth laser scanner and SOFA coordinate spaces ("initial registration" step in fig. 4.15).

Figure 4.16b shows a ground truth laser scan of one of the phantom palpation experiments. Scans were measured using a Nextengine HD laser scanner fixed 20cm from the model scene. Spray chalk contrast agent was applied to the scene prior to each experiment to obtain an accurate surface scan of the palpation tool and phantom. The scanner had sub mm precision and pre-registered grayscale images to 3D meshes in SI units. Ground truth of palpation loads up to 9.8N



(a) Real world view of a silicone phantom palpation experiment. During palpation a load is applied to the phantom using the black palpation tool, which presses into the phantom surface.

(b) 3D laser scan ground truth input to our validation pipeline. We validate our simulation against ground truth by comparing the scanned surface with a simulated surface after simulating the real world palpation load (7.03N in this example).



(c) Our physical simulation method under test. The red object is the simulated position of the palpation tool with 7.03N applied to our phantom MRE model. The phantom model is rendered as a cyan volumetric mesh and has been depressed by the red tool.

(d) Output of our phantom simulator mapped to the ground truth coordinate system. Our phantom simulations agreed closely with the ground truth laser scans. We computed an average 2.14mm L2 registration error with the ground truth scans.

Figure 4.16: Major stages of our MRE image guidance validation pipeline (fig. 4.15) during a palpation experiment. Figure 4.16a is the real world view of the phantom scene. Our goal is to accurately model the displacement of the tool and phantom surface with our simulation method. Figure 4.16b is a ground truth laser scan of the palpation. We compare the ground truth surface with our simulated surface to validate accuracy of our MRE simulation method. Figure 4.16c shows our physical simulation of a palpation experiment. The image guidance pipeline can project guidance landmark positions from our simulation to the real world view during a palpation. Figure 4.16d shows a projection from our simulation to a ground truth scan. The hidden part of the palpation tool has been revealed. For each palpation experiment we validate our method by computing the L2 distance between our simulated phantom and tool surfaces and the set of ground truth laser scans.

was recorded from the Futek micro load cell along with ground truth surface scans of the scene. Much of the phantom and palpation tool geometry were not captured in surface scan ground truth due to the limited spatial range of the NextEngine scanner and self-occlusion. Our data posed realistic challenges to the image guidance pipeline used to validate MRE enriched physical simulations of the scene.

Figure 4.16c shows our patient specific physical simulation method applied to a phantom palpation experiment. The virtual phantom stiffness was determined by pre operative MRE scans of the real phantom. The virtual tool uses load cell data to apply the same force as the real world tool to the simulated phantom. Our simulation depresses the tool into the phantom surface to perform a virtual palpation.

Figure 4.16d shows the video overlay output from our image guidance pipeline. We use the inverse coordinate transform ($\mathbf{T}_n^{-1}$) to map surface landmarks extracted from our physical simulation to greyscale images of the palpation. The simulated geometry of the palpation was also used to validate the accuracy of our simulation (Landmark Placement Delta box in fig. 4.15). We compared the surface of the virtual scene with our ground truth laser scans by measuring surface registration error between the two. Results of the accuracy validation are summarized in table 4.4.

Table 4.4: Phantom simulation error using our data driven method. Our method performed better than the state of the art and was always within acceptable clinical error margins.

	Phantom A			**Phantom B**			**Phantom A/B**		
Load (N)	1.69	3.85	7.81	0.78	4.02	7.03	2.73	4.20	6.42
Error (mm)	3.14	3.24	0.43	2.33	3.86	1.48	1.66	3.10	0.98

Table 4.4 summarizes the validation results from two sets of phantom palpation experiments. On average, registration error in our experiments was 2.14mm which is approximately half that reported in the related work [70, 74] and not clinically significant. Our results indicate MRE is a promising technology for accurate image guidance in liver procedures.

4.6 MRE Stiffness Measurement Validation

Stiffness data must be accurate in order to improve physical simulation accuracy. We verified the accuracy of our data by comparing our MRE model of a silicone phantom with stiffness data measured using a traditional material test of the phantom. We found close agreement between the two measurements, indicating our modeling method is accurate.

MRE Validation Methods

The validation method illustrated in fig. 4.17 compares $\bar{E}$ measured using our MRE modeling method ($\bar{E}_{MRE}$) with $\bar{E}$ measured using traditional material characterization ($\bar{E}_{MS}$). We validated a heterogeneous phantom described by tables 4.2 and 4.3. Details of our validation method follow:

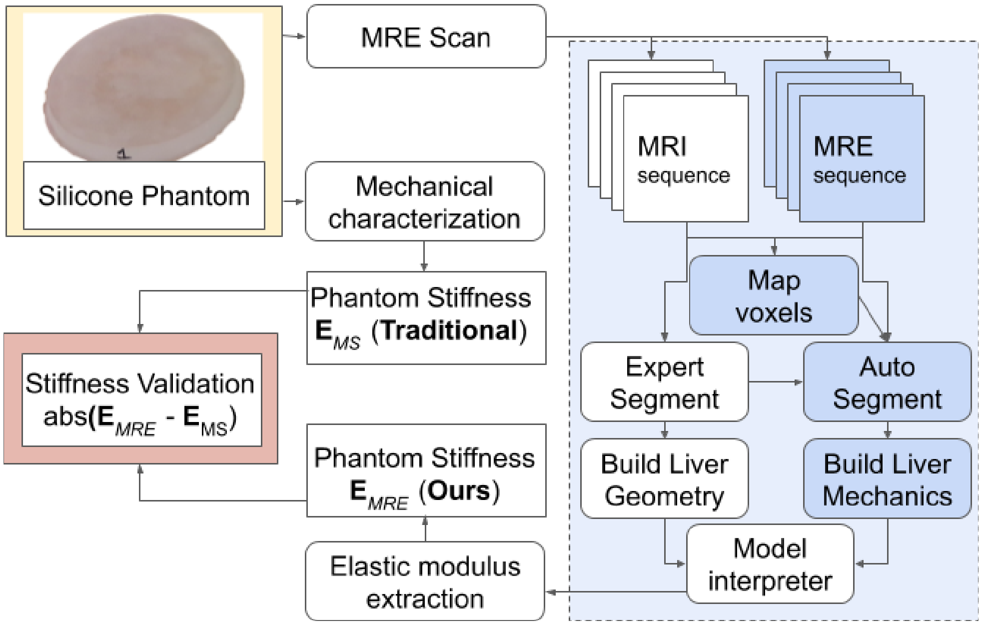


Figure 4.17: MRE stiffness measurement validation. We compare "$\bar{E}_{MRE}$" measured using our modeling method to "$\bar{E}_{MS}$" measured using a traditional material characterization.

First, we took MRE scans of a silicone phantom and ran our patient specific modeling flow denoted by the blue box in fig. 4.17. Next, we manually segmented the scans by each of the silicone compounds in table 4.3. We then extracted $\bar{E}_{MRE}$ stiffness for each compound using the segmentation boundaries and a model interpreter.

In a separate flow, we excised samples from our phantoms and ran a traditional mechanical characterization. We computed $\bar{E}_{MS}$ stiffness using the "Mechanical Characterization" step of fig. 4.17. The traditional characterization flow used a CellScale MicroSquisher. Multiple samples of compounds A and B (table 4.3) were taken from the phantom at various random locations to account for any stiffness variation within the compounds themselves. The MicroSquisher characterized Young's moduli for samples of each cylinder material were then mean averaged to compute $\bar{E}_{MS}$.

Finally, we compared $\bar{E}_{MRE}$ with $\bar{E}_{MS}$ by computing $|\bar{E}_{MRE} - \bar{E}_{MS}|$ for each material in table 4.3 to validate the accuracy of our stiffness modeling method.

Results And Discussion

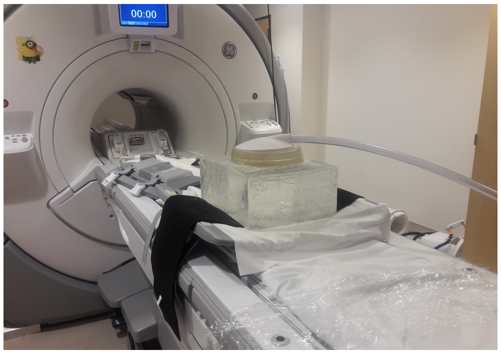
We found close agreement between elastic moduli measured using our MRE method ($\bar{E}_{MRE}$) and those measured using a traditional characterization ($\bar{E}_{MS}$). Figure 4.18 and table 4.5 are a summary of our findings. We discuss our study in detail below.

Table 4.5: Sample MicroSquisher silicone characterization data showing min, $\mathbf{\bar{E}_{MS}}$, and max Youngs modulus for Humimic silicone compounds detailed in table 4.3. The $\bar{E}_{MS}$ modulus for samples of the two silicone compounds agrees within 8% of the MRE $\bar{E}_{MRE}$ moduli; 18.85kPa versus 17.47 for A (852844007437) and 35.07kPa versus 36.81 for B (608729261421).

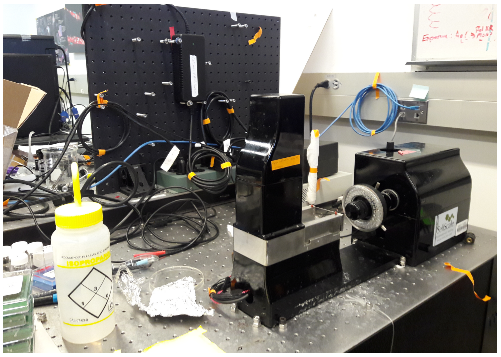
	Cylinder (A)			Cylinder (B)		
Diameter	4924.5um	4989.7	4898.8	5014.3	4931.5	4902.6
Height	4.33mm	4.52mm	4.57	4.16mm	4.6mm	4.41mm
Displacement	701.254um	700.219um	697.634um	709.528um	686.773um	700.736um
Force	25940.8uN	55020.8uN	71292.9uN	69801.9uN	100008uN	71994.6uN
Young's modulus	8.41kPa	**18.163**kPa	24.778kPa	20.725kPa	**35.07**kPa	42.402kPa

Table 4.5 tabulates Young's modulus data from the MicroSquisher experiments. $\bar{E}_{MS}$ is reported in bold for emphasis. In the softer compound, $\bar{E}_{MS}$ agreed to within 8% of $\bar{E}_{MRE}$ (18.85 versus

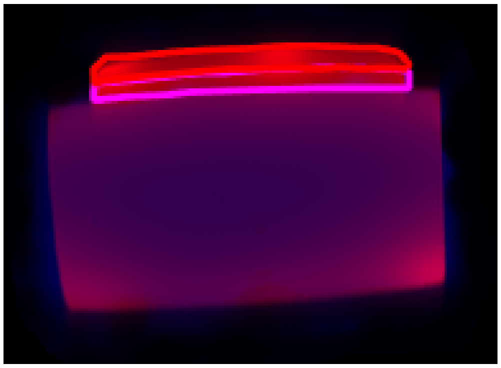
17.47kPa). In the stiffer compound, agreement was within 5% (35.07 versus 36.81kPa). The data indicated our traditional characterization method was accurate with imprecise outliers, which we report in table 4.5 either side of $\bar{E}_{MS}$ for both silicone compounds in the phantom under test.



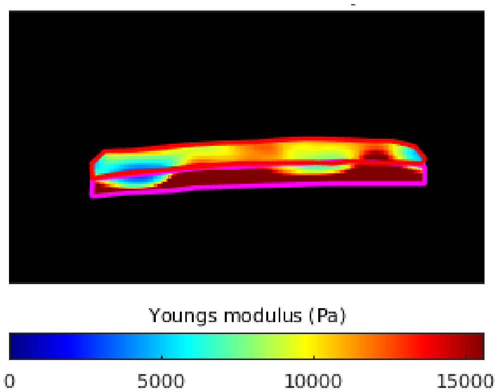
(a) Silicone Phantom MRE arrangement. The Cylindrical Phantom of interest is mounted on an oblong torso chassis in order that the abdominal MRE protocol used for liver scans may be run on it. The MRE actuator paddle is seen mounted on the cylindrical phantom



(b) Silicone mechanical characterization setup. 3mm tubular samples of the phantom are subjected to stress relaxation characterization by the MircoSquisher.



(c) Combined image of phantom MRI (red) and overlaid phantom MRE (blue). The contrast of the coronal scans has been enhanced to make segmentation of the different silicone compounds in table 4.3 easier. MRE masks used for computing the Young's modulus in the cylindrical phantom are depicted by red and magenta lines respectively



(d) Example segmented MRE scan. MRE masks used for computing the Young's modulus are depicted by red and magenta lines respectively. Our MRE protocol exhibited artifacts at phantom tissue boarders, which made segmentation of tissue types and averaging of the elastic modulus necessary to compensate.

Figure 4.18: Our MRE validation method. Phantom MRE stiffness measurements experiments for $\bar{E}_{MRE}$ and $\bar{E}_{MS}$ are depicted in fig. 4.18a and fig. 4.18b respectively. Figure 4.18c illustrates how we adapted our modeling flow to isolate $\bar{E}_{M}RE$ for silicone compounds of different stiffness. Figure 4.18d is an example segmented stiffness map of one of our silicone phantoms.

Figure 4.18a shows our arrangement for taking the phantom scans we used to compute $\bar{E}_{MRE}$ with our patient specific modeling method. We used a GE Medical Systems Discovery MR750 with 3.0T field strength set to its abdominal protocol at 50% amplitude. The protocol was identical to that used in our previous studies in sections 4.3 to 4.5. In order to use the abdominal protocol, the phantom was mounted to an oblong silicone torso phantom chassis and then interfaced to a MRE paddle actuator.

Figure 4.18b shows our traditional material characterization experiment. 3mm diameter tubular samples were taken from random locations of the phantom. A linear elastic modulus of the samples was characterized using the dry compression test series of the MicroSquisher. The Poisson ratio (v) of each sample was then computed. We used the reported compression strain along the straight edge of the samples and the diameter change at the center of the samples measured by the MicroSquisher camera when computing v.

Figure 4.18c shows compound segmentation of the heterogeneous silicone phantom in MRI and MRE scans of the transverse imaging plane. The MRI image has been color stained red and the corresponding MRE has been overlaid using the "map voxels" step from our liver modeling method. We used visible contrast boundaries between silicone compounds to perform manual segmentation of the MRI scans. The MRE scan, color stained blue, was segmented by silicone compound using the previously described MRI boundaries and the "map voxels" subroutine. Silicone compound boundaries are indicated by magenta and red outlines.

Figure 4.18d illustrates a MRE stiffness map of a heterogeneous cylindrical phantom along the transverse imaging plane. $\bar{E}_{MRE}$ was computed for each compound as the average stiffness of all voxels for that compound, where E for each voxel is found using eq. (4.1). We computed two distinct $\bar{E}_{MRE}$ Young's moduli for compounds A and B in table 4.3 of 17.47 and 36.81kPa respectively. The softer modulus of the phantom was between the fourth and fifth standard deviation of non diseased liver and the stiffer modulus was between the third and fourth standard deviation of stage 4 liver fibrosis [75]. Overall, MRE indicated the phantom stiffness was representative of diseased liver tissue.

4.7 Conclusion

We introduced a novel patient specific physical simulation approach for accurate surgical landmark tracking that incorporates Magnetic Resonance Elastograms (MRE). We used several studies to motivate, demonstrate, and validate our patient specific physical simulation method. Our preliminary results indicate that liver stiffness may vary across patients and within patients and that tracking surgical landmarks with our physical simulation is more accurate than conventional methods that do not directly measure patient specific biomechanical stiffness.

First, our statistical cohort analysis of patient MRE scans motivated the use of patient specific stiffness in simulation. The study found large variation in liver stiffness between patients and within patient livers that was significant when compared with the atlas stiffness used in traditional simulators. Second, our study on surgical landmark placement differences compared our MRE method with a traditional simulation of a liver retraction procedure on our patient cohort. We found clinically significant differences in accuracy between the two methods which motivated further validation of our MRE method in an image guidance pipeline. Third, our silicone palpation validation used an image guidance pipeline instrumented for recording ground truth to validate the accuracy of our proposed patient specific physical simulation method. Our MRE simulation pipeline was able to track a palpation of silicone phantoms with double the accuracy on average as compared with prior work. Finally, our evaluation of MRE stiffness measurement accuracy used a heterogeneous silicone liver phantom to demonstrate the accuracy of MRE based biomechanical models, which we previously found to be important in accurate landmark placement. We found close agreement between MRE scans of our phantom and a conventional material characterization using a material tester.

Our results are promising and motivate further animal model studies, which are necessary to better understand the clinical significance of MRE enriched physical simulation in surgical image guidance applications.

This chapter, in full, is a reprint of submitted material as it may appear in Medical Image Analysis (MIA), Barrow, Michael; Chao, Alice; Fowler, Kathryn; He, Qizhi; Ramamoorthy, Sonia and Kastner, Ryan, 2020. There are small changes in format and phrasing as a chapter within this larger paper. The dissertation author was the primary investigator and author of this paper.

Chapter 5

Empirical Method For Patient Specific Image Guidance

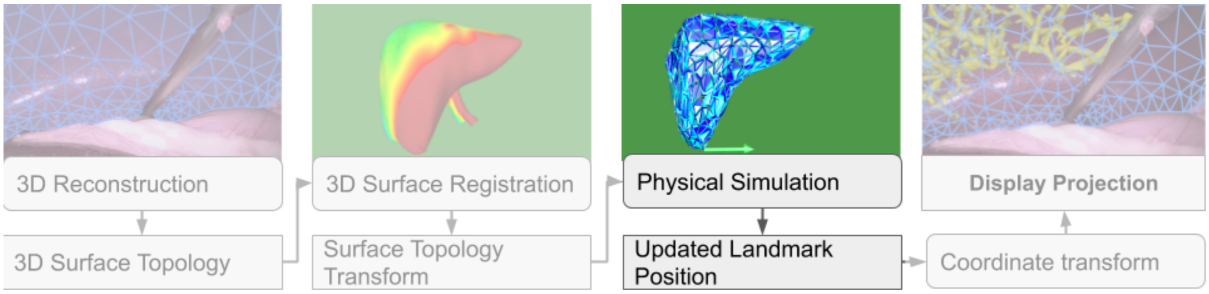


Figure 5.1: Our Empirical physical simulation method improves efficiency of the online landmark placement step in a CAIG pipeline.

Recall the four major online steps required for accurate CAIG outlined in chapter 1. In this chapter we describe a "Empirical Method For Patient Specific Image Guidance". This method applies to the physical simulation step. Figure 5.1 highlights the simulation step in the CAIG pipeline to illustrate where our contribution is useful within a larger CAIG system. Our experiment data driven method improves the accuracy of landmark placement without impacting the speed of the simulation step. It therefore improves efficiency of CAIG.

Introduction

It remains challenging to register preoperative images to the surgical site. Due to the constant and significant change of the position and shape of the liver, manually mapping preoperative image(s) onto the surgical scene is time consuming and error prone [214], resulting in more conservative surgical decisions.

Computer Aided Image Guidance (CAIG) fuses preoperative scans with intraoperative images to provide more detailed information about the surgical site. For example, Augmented Reality Image Guidance (ARIG) merges the preoperative data directly into the surgeon's view and Virtual Reality Image Guidance (VRIG) provides different views of the preoperative and intraoperative data, e.g., as a manipulable 3D model [44]. Accurate real-time CAIG is a valuable and important surgical tool that lead to more precise surgical procedures.

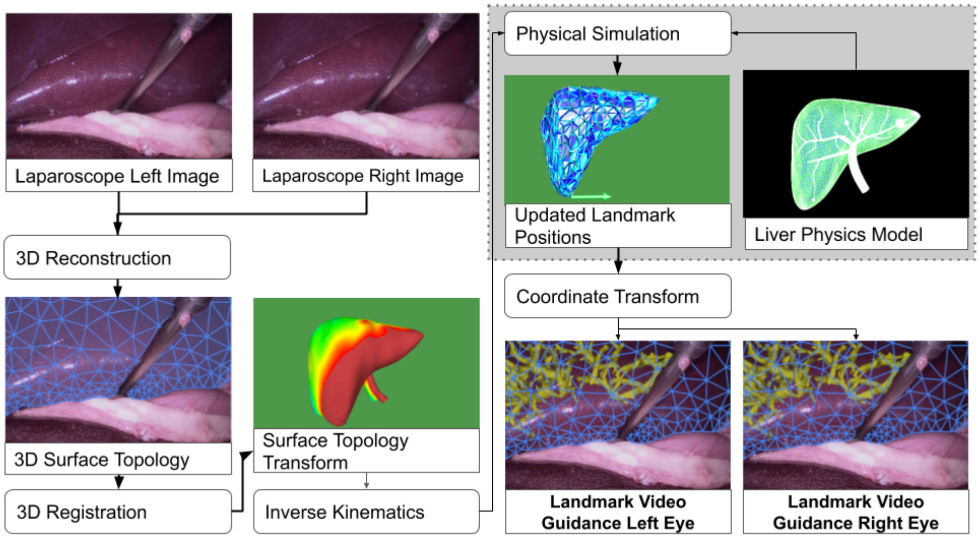


Figure 5.2: Block diagram of an CAIG pipeline. For each video frame: Firstly, the 2D laparoscope video is converted to 3D. Secondly, registration maps the visible portion of liver to a complete 3D surface model. Thirdly, changes on the 3D surface are propagated inside the liver using a physical simulation. Finally, updated guidance landmark positions are displayed on a modified laparoscope video feed. This work focuses on accurate physical simulation (gray box) which is key to accurate IG.

The state of the art in computer assisted image guidance for hepatic surgery does not use patient specific biomechanical models to position surgical landmarks in either computer aided preoperative planning tools [25, 215, 216, 217, 218] or AR image guidance [70, 69, 219, 220]. This lack of accurate liver behavior is a well known drawback of these systems, and leads to poor results [221]. Our work in chapter 4 using patient Magnetic Resonance Elastograms (MRE) indicates: **Higher accuracy in CAIG is possible with patient specific biomechanical modeling**

Method

In this work, we propose a machine learning enhanced data-driven approach for patient specific biomechanical simulation for computer aided image guidance. Our work replaces the gray box in fig. 5.2. The manifold-learning method developed in [222] models liver tissues by using a few measurement data collected experimentally at the the start of the liver procedure. The proposed approach allows to on the fly construct a locally convex material manifold on the liver measurement data usually described by strain-stress relationships, which is integrated with the physical models to perform mechanics based simulation. As a result, the proposed data-driven physical simulation circumvents the necessity of using phenomenological constitutive models that are typically non patient specific and have difficulty describing complex biological materials. This provides real time placement of surgical landmarks under a novel patient specific liver physics model. The proposed approach has two key advantages over prior work.

Firstly: ***Liver model accuracy can be improved***. Prior MRE models only considered linear elastic materials, whereas liver tissue is commonly considered a hyperelastic material. This model error reduced the power of our patient specific image guidance simulations. However, the proposed data-driven approach considers more complex material behaviors.

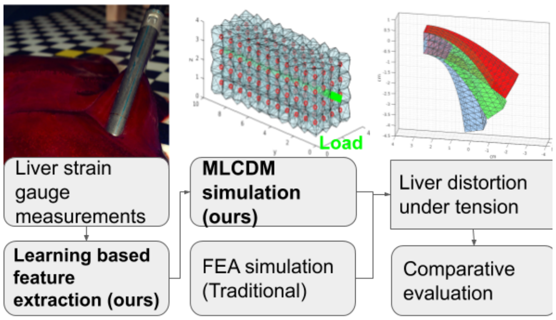
Secondly: ***Clinical significance can be magnified***. Early work relied on Magnetic Resonance Elastography (MRE), which is an expensive and exotic imaging modality. Although laparoscopic surgery is routine, resource limited institutes could not benefit from our MRE technique. Conversely, our proposed method can build patient specific biomechanical models using cheap strain gauges to parameterize the Manifold Learning Convexity Data-driven Modeling (MLCDM) framework [222].

Results

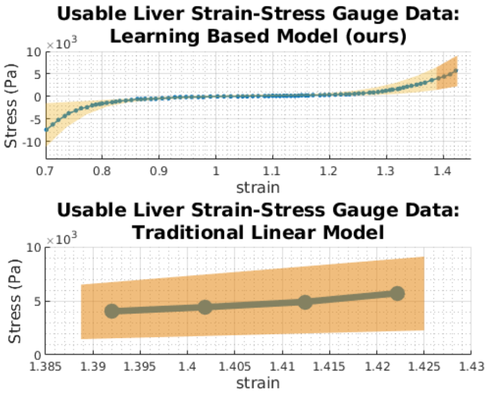
A liver retraction load simulated using our method was compared with a traditional simulation. The experiment is depicted in fig. 5.3. We applied the MLCDM technique to a series of strain-stress gauge measurements of liver tissue taken from the literature [65]. Our preliminary results show close agreement (< 0.1mm) between our data-driven method and a traditional numerical approach to image guidance landmark placement. In addition, fig. 5.3b shows our method is able to model 25 times more experimental observations than the traditional method.

In our next steps, we will evaluate MLCDM in vivo. Although we found agreement with traditional methods in the linear strain-stress region of liver data, MLCDM can better model clinically significant non-linear mechanics that are important to accurate image guidance landmark placement. The MLCDM technique is very simple to integrate with the surgical work flow. Unlike MRE, no complex segmentation, mapping or interpolation of stiffness data is required. Measurements can directly be used by the MLCDM physical simulator since it extracts features on the fly and performs numerical optimization directly.

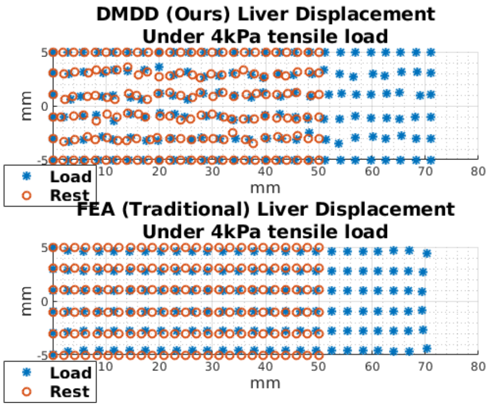
This chapter, in full, is a reprint of submitted material as it appears in the third Black in AI Workshop (BAI), Barrow, Michael; He Qizhi, and Kastner, Ryan, 2019. There are changes in format and phrasing as a chapter within this larger paper. The dissertation author was the primary investigator and author of this paper.



(a) Learning based IG simulation method



(b) Liver load response data



(c) Ours vs traditional simulation

Figure 5.3: Fig 5.3a shows an evaluation of the proposed patient specific simulation. Several measurements are taken with a liver strain gauge before running our IG method. Much more of this liver data can be used to position surgical guidance landmarks in our learning based MLCDM simulation compared to traditional methods as shown in fig. 5.3b. In fig. 5.3c we compare our IG approach to the traditional. We find good agreement between our learning based IG method and traditional FEA when simulating a liver retraction, a common step to "hoist" the liver during surgery.

Chapter 6

Data Driven Efficiency Optimization for Portable Hardware Accelerated Image Guidance Platforms

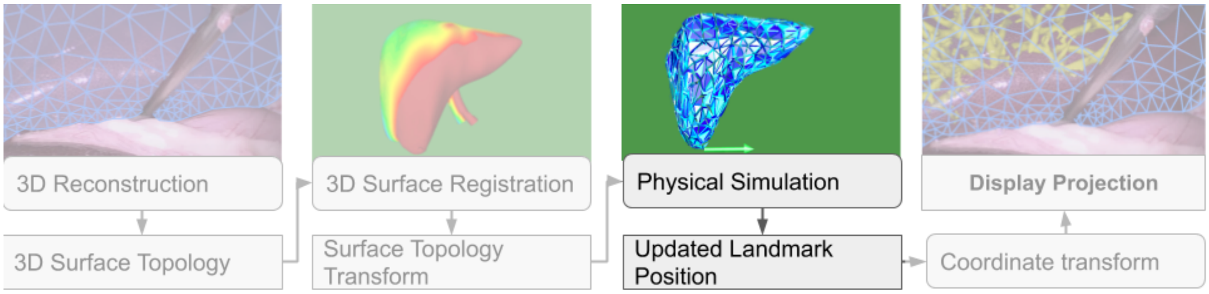


Figure 6.1: Our optimization framework improves efficiency of hardware accelerated CAIG platforms. The framework finds the minimal power dissipation of a CAIG platform that meets an accuracy requirement.

Recall the four major online steps required for accurate CAIG outlined in chapter 1. In this chapter we describe "Data Driven Efficiency Optimization for Portable Hardware Accelerated Image Guidance Platforms". This method is demonstrated on the physical simulation step. Figure 6.1 highlights the simulation step in the CAIG pipeline to illustrate where our contribution is demonstrated within a larger CAIG system.

6.1 Introduction

High precision surgical robots are fast becoming common place in the operating room. While these robots promise sub millimeter surgical precision, this has been locked behind a wall of traditional video imaging systems where visual estimation is the only precision available. Much recent effort has been invested in Computer Aided Image Guidance (CAIG) techniques that fuse static preoperative images with the dynamic video feed. However, focus has been somewhat narrow and questions of practicality have been overlooked. Although the common assumption is that high end accelerators can be used for CAIG, this difficult to achieve in many routine procedures. Consider the photograph taken during a lower anterior resection.

Figure 6.2: Photograph during Robotic Assisted Laparoscopic Lower Anterior Resection. This case required resection and re-connection of the large intestine using a transanal circular stapler. A Davinci Xi was used to perform the resection and insertion of one stapler face, the other was manually inserted. The challenge of this procedure is mating the stapler faces within the patient in a "space dock" maneuver. The stapler operator has no visibility and must rely on her colleague to guide this part of the procedure.

Three surgeons are connecting a resected large intestine using a surgical robot and a manually operated transanal stapler. Although the surgeon remotely operating the robot has a free work area where a large GPU accelerated machine can provide CAIG, the stapler operator has a highly restrictive work area. Indeed, she must rely on audio guidance provided by a colleague with a view from the robotic laparoscope. A head mounted CAIG platform is essential in such situations.

We focus on the lightly explored but highly critical research question of how to optimize the efficiency of physics acceleration for head mounted CAIG. Physical simulation is a computationally intensive essential step for accurately positioning image guidance landmarks. It must be accurate and fast enough for the real time clinical application and a portable platform adds additional constraints. Firstly are patient safety concerns. The device must remain sterile and guidance cannot be interrupted. This means trailing power chords, battery changes and wireless displays are unacceptable. Secondly, the simulation must be energy efficient enough that the platform battery can last the whole procedure (2-12+ hours). Thirdly, the platform cannot draw more power than would cause the care giver discomfort and thereby jeopardize patient safety. Finally, the platform should be as unobtrusive as possible. For example, a platform incorporating a head mounted display connected to a backpack device is acceptable, however it is clear from fig. 6.2 that this would restrict movement and is undesirable.

Because of the conflicting requirements of performance and efficiency, it is not clear what the best way is to provide CAIG to surgeons physically constrained by a procedure. Our prior experience with the state of the art SOFA CAIG physical simulator [210] could not achieve acceptable performance with CPU only simulation in liver procedures and hardware acceleration of physical simulation is needed. Our research question was, for a given surgical case, how unobtrusive could a system could be? for example if a procedure on a 1000cm^3 liver required $<$ 5mm accuracy for 5 hours, would a backpack PC based accelerator have a battery capable of physical simulation for the procedure? Or would a low power head mounted accelerator be accurate enough? How does the accelerator type affect the answer? GPUs are generally assumed to be less power efficient than FPGAs but are easier to program and are supported by SOFA. FPGAs are acknowledged as being more flexible than GPUs and recent tooling innovations such as OpenCL have closed the programability gap.

In this paper we develop a framework to address these questions. The rest of the paper is organized as follows. Section 6.2 describes the related work. section 6.3 outlines physical simulation and how we map it to a hardware accelerator for evaluation. Section 6.4 details our experiments to measure the efficiency of PC and embedded GPU and FPGA accelerators in physical simulation of a range of liver cases. Section 6.5 reports our findings and demonstrates how our methods can be applied to find optimal accelerator schemes for a given CAIG platforms, power and energy constraints. Section 6.6 draws our conclusions.

6.2 Related Work

CAIG applications have traditionally been limited by the highly challenging computer vision problems of surgical scenes coupled with stringent clinical accuracy requirements. All augmented reality (AR) CAIG requires a "registration" step where the real world co-ordinate system is mapped to a virtual one containing guidance cues. Accurate "rigid" registration techniques such as Iterative Closest Point [223] have been proposed which perform a singular decomposition of the source image co-ordinate system to find a global transform mapping the co-ordinate systems. Although the number of registration points is often large, many real time variants have been proposed at the same time hardware capabilities have improved [224, 225, 226]. Such algorithms are useful for CAIG in procedures where important structures in the surgical scene can be assumed “rigid” such as nephrectomy and osteopathy [101, 227]. However, the rigid assumption often breaks down in abdominal surgeries and many local warp functions must be computed and optimized for reasonable registration which is much more computationally demanding.

Since optimizing the warp functions has a lot of parallel work, much effort has been made to apply hardware acceleration to CAIG. Recently real time non rigid registration has been demonstrated on GPU [167] and FPGA [163]. However CAIG for abdominal procedures requires an additional physical simulation step to compute the warping of guidance cues that are beneath a visible surface in the surgical scene. CAIG Physical simulation techniques are typically limited to linearized models of tissue owing to the order of magnitude computuational overhead of more descriptive non-linear models. Although non-linear simulations have been proposed, video rate performance requires unreasonably low resolution models compared with linear approximations even when accelerated on GPU [204, 121]. Additionally validating non-linear biomechanical models in-vivo is much more challenging and makes this trade-off more unattractive [228].

Using linear biomechanical models, CAIG physics simulation based blood vessel avoidance has been demonstrated with an estimated 5mm in-vivo accuracy in liver procedures [70]. The authors used the SOFA surgical simulation framework [210] which implements hardware accelerated physical simulation [229] to achieve the necessary video rate of physical simulation. This is possible since FEM and other discrete methods of simulation have much parallel work that benefits from hardware acceleration and CAIG frameworks such as SOFA have arisen that package various parallel physics simulators [230, 121, 203].

Despite this progress, CAIG is difficult to harness thanks to unwieldy hardware accelerator platforms. Traditionally researchers have paid little attention to the power and thermal footprint of CAIG platforms since clinical margins and video frame rates are usually the prime focus. Ignoring the physical requirements of CAIG platforms limits the clinical relevance of recent advances. In most cases devices are too bulky for use with anything besides remote controlled robotic tools and or create operating room hazards by relying on trailing cables for power [231, 232, 233]. Getting CAIG to clinicians requires simulation on unobtrusive personal CAIG platforms to avoid disrupting their work flow.

Recent developments in portable augmented reality hardware present a great opportunity to address this need. Portable AR was first applied to non-medical image guidance decades ago. However, the first devices were cumbersome. Early examples include Feiner [234] and Höllerer [235] who both used laptop technology to create 40 lb prototypes designed to be strapped to the user's back and head. These "backpack" prototypes touted sufficient computational power to run an AR programs that displayed helpful information about the user's surroundings as they explored an outdoor area. Since then, smaller, lighter, and commercially available products like the Google Glass and smart phones have been successfully used in a variety of AR applications [236, 237] including CAIG.

Although mobile devices feature the same types of GPU and FPGA hardware accelerator long used in CAIG, the accelerators have less computational capability due to reduced power sources and less capable thermal solutions. Although much research effort has been made in understanding mobile registration acceleration [238, 239], little literature explores mobile physical simulation for CAIG.

It is therefore important to better understand how the clinical demands of an abdominal procedure can be most optimally mapped to a CAIG physical simulation accelerator hardware so that clinicians can apply this promising technique in as many surgical cases as possible.

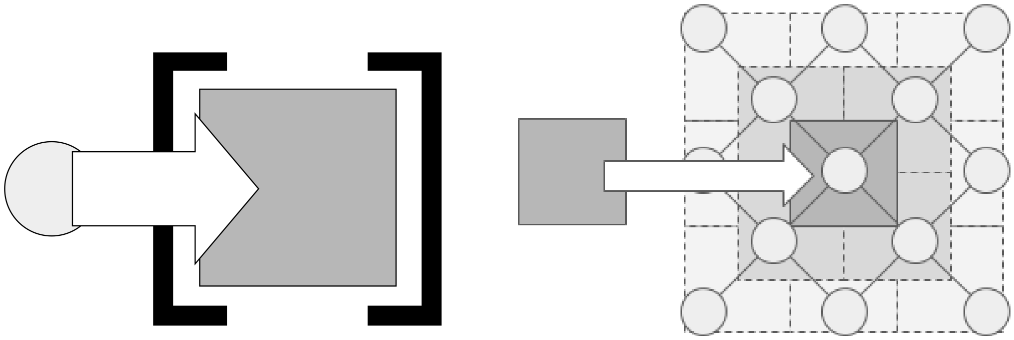
6.3 Method

We perform a series of experiments on several candidate hardware accelerator platforms. Our goal is to understand how well suited each of the platforms under test are to portable CAIG applications. In particular we focus on real time physical simulation of the liver which is the essential CAIG step for accurately positioning IG landmarks such as vessels and tumors. Specifically, the problem to solve on hardware is: Video rate physical simulation of a liver in a mobile power envelope. Our method measures the power consumption of hardware accelerators running our physical simulation model (described in section 6.3.1) as power traces. These traces are combined with additional information about simulated CAIG procedures and the hardware to produce temperature and battery performance plots.

6.3.1 Accelerator Physics Simulation Mapping

Although various order reduction and refinement strategies exist to reduce simulation problem sizes and improve accuracy, in-vivo verification of these sophisticated strategies remains challenging. Furthermore computational savings are lost when significant changes in the liver topology occur and this can complicate choosing an optimized model. We therefore assume simple sparse, heterogeneous, and isotropic conditions to create a liver model that is well suited for surgical image guidance applications and can be efficiently mapped to an accelerator.

We assume a mesh-free simulation approach such as that proposed by Faure et al. [188] for the SOFA surgical simulator. The key advantage in this method is that re-meshing is not required when the model undergoes topological changes which results in consistent performance. We then perform even sampling of the liver geometry in $\mathbb{R}^3$ and assume a simple "domain of influence" scheme similar to that used by Dolbow et al. [240]: only neighboring nodes within a preset cubed radius from the center node are included in the local stiffness matrix of the center node. The resultant global stiffness matrix becomes a block diagonal matrix with each block matrix representing the the local stiffness matrix of one node's domain of influence in a similar fashion to Nealen's approach [241]. The "domain of influence" scheme is illustrated in an example in fig. 6.3.



(a) Local stiffness matrix for a node

(b) example domain of influence

Figure 6.3: Figure 6.3a represents the stiffness matrix of a single node with no neighbors and fig. 6.3b illustrates a local stiffness matrix: a node's domain of influence in 2D. Only neighbors within a square radius of influence are considered to affect the single central node. The resultant local stiffness matrix ignores nodes outside the square radius which can cause large oscillations about the $\delta\dot{\mathbf{q}}$ in the CG descent.

The "domain of influence" strategy allows us to pare the local stiffness matrix down to a relatively small size and we can further reduce the problem size by multiplying nodally which allows us to reduce the size of the global stiffness matrix in the following way. Recall the canonical semi-discrete formulation given in eq. (4.2) and its implicit Euler mappings given in eqs. (4.4) to (4.6). Since in our scheme, the global stiffness matrix $\mathbf{K}$ has been defined as a block diagonal matrix, multiplying each local stiffness matrix A_k with its corresponding portion of the $\delta\dot{\mathbf{q}}$ vector $\mathbf{x_k}$ is equivalent to $A\mathbf{x}$ for all $1 \leq k \leq N$ with N being the total number of nodes. Compiling the $A_k\mathbf{x_k}$ vectors into one vector in the same order they were multiplied would result in $A\mathbf{x}$. Thus we can rearrange all of the local stiffness matrices that comprise the global stiffness matrix and their corresponding portions of the $\delta\dot{\mathbf{q}}$ vector.

$$A\mathbf{x} = \underbrace{\begin{bmatrix} A_1 & \dots & & & \\ \vdots & A_2 & & & \\ & & A_3 & & \\ & & & \ddots & \vdots \\ & & & \dots & A_N \end{bmatrix}}_{Nn \times Nn} \underbrace{\begin{bmatrix} \mathbf{x_1} \\ \mathbf{x_2} \\ \mathbf{x_3} \\ \vdots \\ \mathbf{x_N} \end{bmatrix}}_{1 \times Nn} \tag{6.1}$$

where each A_k for $1 \leq k \leq N$ and $k \in \mathbb{N}$ is of dimension $n \times n$, all other entries of A are 0, and N represents the total amount of local stiffness matrices in the global stiffness matrix. We can modify A and $\mathbf{x}$:

$$A_{mod} = \underbrace{\begin{bmatrix} A_1 \\ A_2 \\ A_3 \\ \dots \\ A_N \end{bmatrix}}_{n \times Nn} \tag{6.2}$$

$$\mathbf{x_{mod}} = \underbrace{\begin{bmatrix} \mathbf{x_1} & \mathbf{x_2} & \mathbf{x_3} & \dots & \mathbf{x_n} \end{bmatrix}}_{N \times n} \tag{6.3}$$

$$A_{mod}\mathbf{x_{mod}} = \underbrace{\begin{bmatrix} A_1\mathbf{x_1} & \dots & & & \\ \vdots & A_2\mathbf{x_2} & & & \\ & & A_3\mathbf{x_3} & & \\ & & & \ddots & \vdots \\ & & & \dots & A_N\mathbf{x_N} \end{bmatrix}}_{N \times Nn} \tag{6.4}$$

The product of $A_{mod}\mathbf{x_{mod}}$ results in a $N \times Nn$ matrix whose diagonal can be partitioned into $1 \times n$ matrices $A_k\mathbf{x_k}$. Extrapolating this diagonal reveals:

$$\underbrace{\begin{bmatrix} A_1\mathbf{x_1} \\ A_2\mathbf{x_2} \\ A_3\mathbf{x_3} \\ \vdots \\ A_N\mathbf{x_n} \end{bmatrix}}_{1\times Nn} = A\mathbf{x} \tag{6.5}$$

Modifying A and $\mathbf{x}$ so that all trivial multiplications with 0 are eliminated allows the product to be computed more efficiently. Since memory operations are high in latency, power hungry, and contribute to the bottleneck created by the CG linear solver, reducing memory traffic is in line with our goal for efficient physical simulation in embedded devices. This is referred to as an explicit formulation in the literature [241] and is applied to hardware accelerated physics in CAIG [242, 243]. By adding the additional constraint of uniform internal nodes we exploit further parallelism in a batch vector multiplication scheme. We use this method as our CAIG physics simulation model in the remainder of the paper.

6.3.2 Simulation Problem Size

We defined the simulation problem size by considering patient liver volume and the simulation accuracy impact of discretizing this volume. Maximum and minimum volumes were taken from Abadalla's recent study [8]. A summary of the range of volumes considered is given in fig. 6.4.

Acceptable discretization resolutions were chosen using a simulation convergence study with a Euler-Bernolli theoretical model. A cantilever beam was modeled using the Euler-Bernolli equation eq. (6.6) and a second beam built using the scheme outlined in section 6.3.1 was simulated to convergence with the theoretical model.

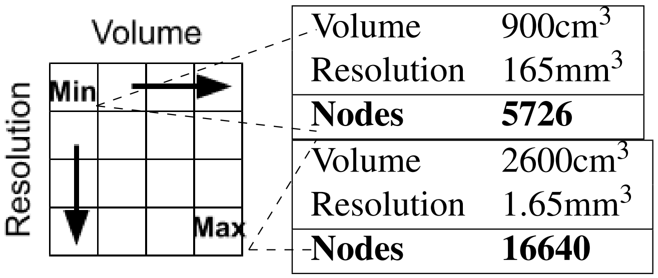


Figure 6.4: Summary of Liver volumes and resolutions studied. 100 data points per device were gathered between the limits of volume and resolution to study the trend of physics accelerator performance. Volume and resolution were swept across 10 discrete values

$$w(x) = \frac{f_{load} x^2 (6L^2 - 4Lx + x^2)}{24EI} \tag{6.6}$$

where $w(x)$ is displacement of the central beam axis in the load direction, E is the elastic modulus, L is the long beam edge and I is the second moment of hyperrectangle cross sectional area. Details of the simulation parameters for the theoretical and mesh free model are provided in table 6.1.

Table 6.1: Convergence simulation study hyperrectangle parameters

L(cm)	W/H(cm)	rho	E(kPa)
5	1	997	12

Our study used a surgical simulation validation strategy employed by Plantefeve et al. to tune liver simulations for image guidance [70]. The method applies a "Pneumoperitoneum" load along the beam which simulates pressure on the liver caused by insuflation of the abdomen at the start of a procedure. Figure 6.5 shows the impact on accuracy of sweeping resolution of the volume discretization from a maximum resolution of 1.65mm^3 to a minimum of 165mm^3. We did not reduce discrete volumes below the voxel size of standard MR images since simulation objects are typically constructed using volumetric MR data. We did not increase discrete volumes beyond 165mm^3 since the convergence error of 8mm would not be useful in many abdominal procedures [44].

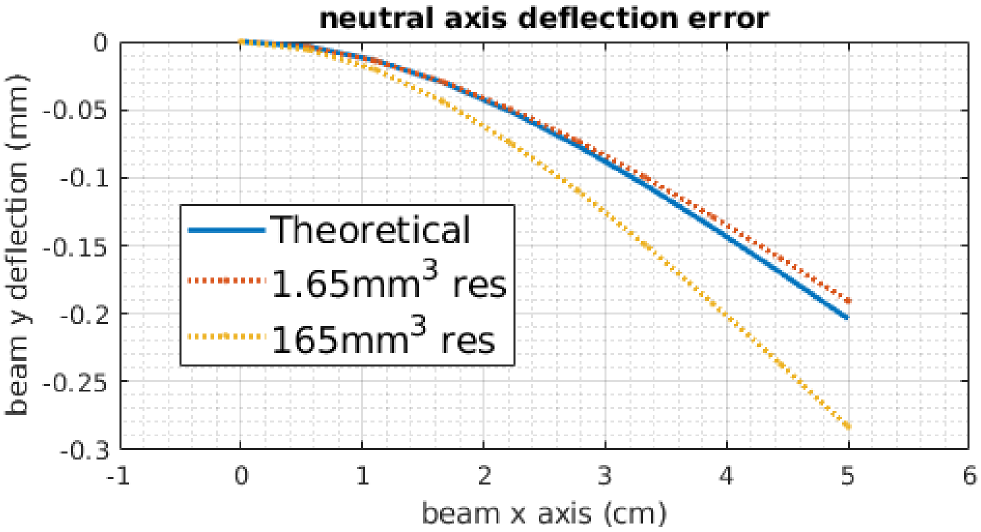


Figure 6.5: Liver simulation convergence with a Bernolli Euler theoretical model. Errors varied between 1-8mm for a fixed beam length of 50mm. Simulation resolution was swept between 1.65-165mm^3.

6.4 Experiment

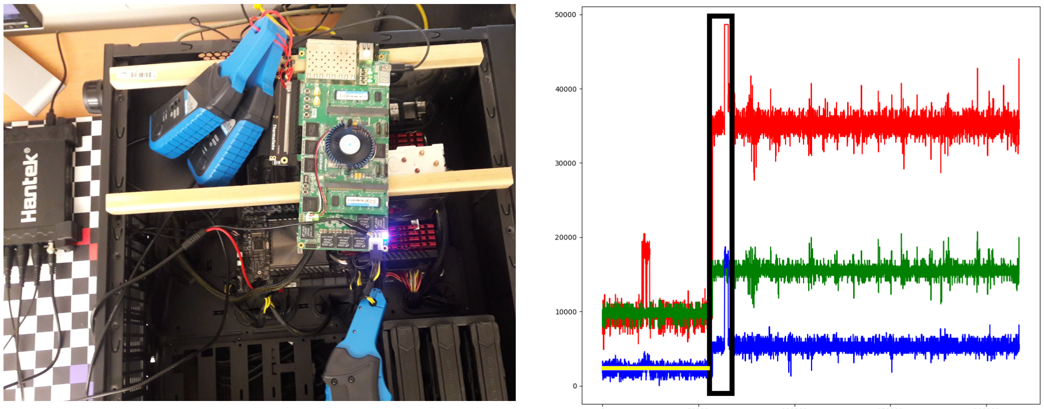
We compared PCIe and embedded accelerator performance in solving liver physical simulation problems. The goal was to understand the interplay between liver size, simulator resolution, hardware temperature and platform battery life. To do this we ran linear solver acceleration benchmarks comprised of optimized batch vector multiplication code for the simulation procedure described in section 6.3.1 to evaluate several CAIG platform specifications outlined in table 6.2.

Our linear solver benchmarking code could be configured to a variety of liver sizes and resolutions, allowing us to sweep those parameters of interest. It is difficult to compare the accelerators to each other due to different architectures and noise on the host platforms. Our benchmarking framework solved this problem by using device specific optimized vendor provided acceleration kernels to accelerate the computational bottleneck and a common high level core for all devices to sweep the same parameters of interest on all platforms. As the benchmarking core

Table 6.2: Candidate platforms in our study of CAIG physics acceleration. Head mounted platforms are development boards with display ports strapped to the clinicians head directly. We assume smart watch batteries would be required for unobtrusive head worn devices. ‡For back mounted platforms we assume the PC PCIe cards would be driven by a laptop-like platform strapped to the clinicians back and use Abadalla's power data to model such a platform [8]. We further assume a laptop battery would be acceptable for the back mounted devices.

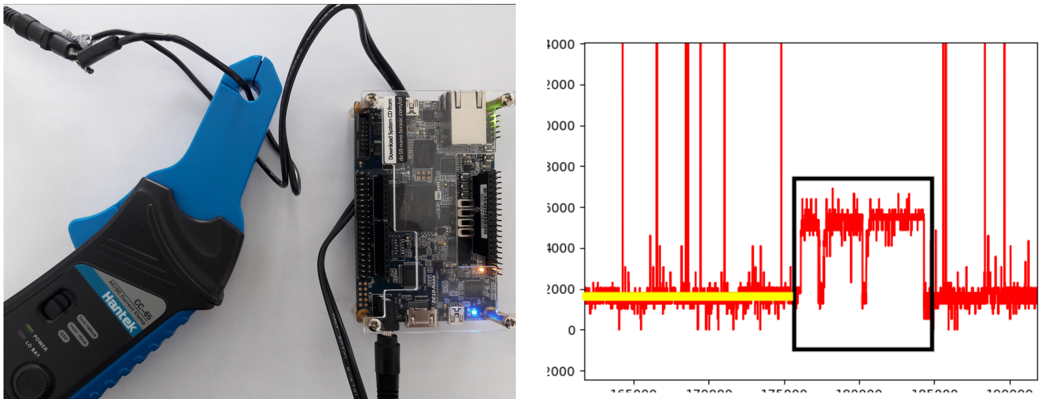
Accelerator	**Worn**	**Voltage (V)**	**Idle P (W)**	**Battery (mAh)**
DE5 FPGA	Back	12	30‡	7700
K40 GPU	Back	12	30‡	7700
DE10 FPGA	Head	5	5.5	1000
TX2 GPU	Head	9	2.8	1000

swept simulation parameters, power traces were obtained using contact free current sensors and the core logged ancillary information from software to validate kernel run time. We addressed platform noise by using minimal software stacks on embedded devices and by instrumenting the accelerator power buses directly for PCIe cards. Instrumenting the PCIe cards allowed us to directly measure accelerator physics simulation kernel power. For embedded platforms we isolated the kernel power by subtracting mean static platform power from accelerator power traces. Each device was instrumented with one or more Hantek CC650 AC/DC current clamp sensors as shown in figs. 6.6a and 6.6c.



(a) Bench marking of back mounted FPGA device

(b) back mounted device benchmark power trace

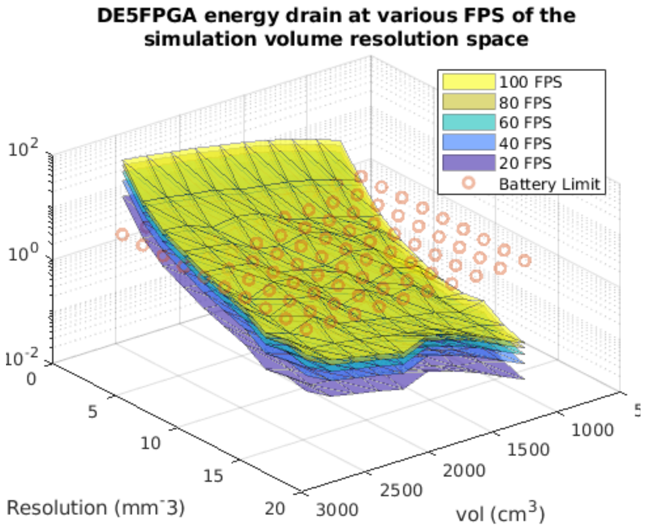


(c) Bench marking of Head mounted FPGA platform

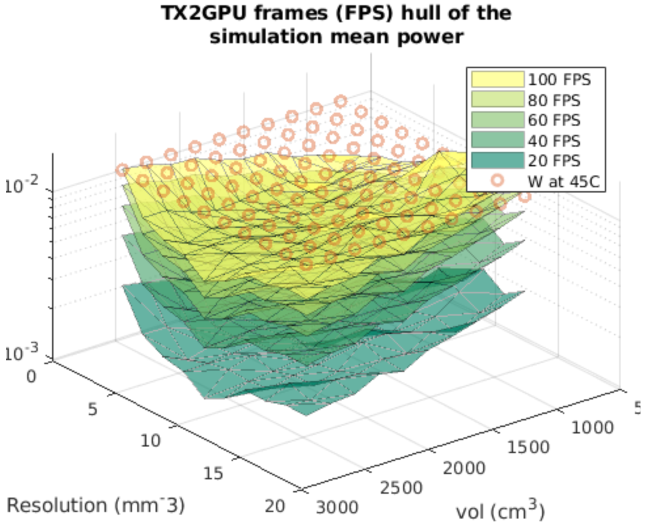
(d) head mounted platform benchmark power trace

Figure 6.6: Examples of our benchmarking method applied to back mounted figs. 6.6a and 6.6b and head mounted figs. 6.6c and 6.6d platforms. Our benchmark modeled the workload of an efficient hardware accelerated Implicit Euler physical simulation method described in section 6.3.1. Our benchmark used fig. 6.4 to sweep liver size and simulation resolution when gathering power traces.

Data from the clamps was measured using a four channel Hantek 6074BE oscilloscope. Logs from the scope were analyzed by an in house tool. The tool estimates battery and thermal performance of an CAIG platform by using key hardware metrics in its thermal model and computing battery life. The model uses 100 power traces from each device to compute the reported trends. A three dimensional visualization of the computed battery and power trends is given in fig. 6.7.



(a) DE5 FPGA Energy Trend

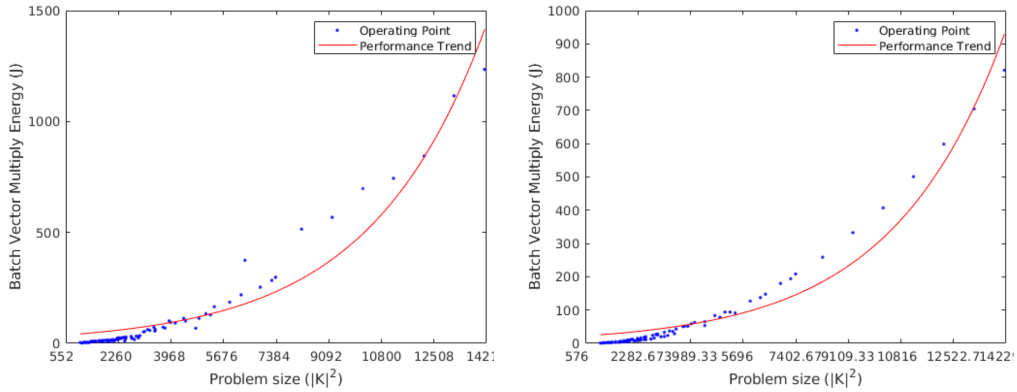


(b) TX2 GPU Power Trend

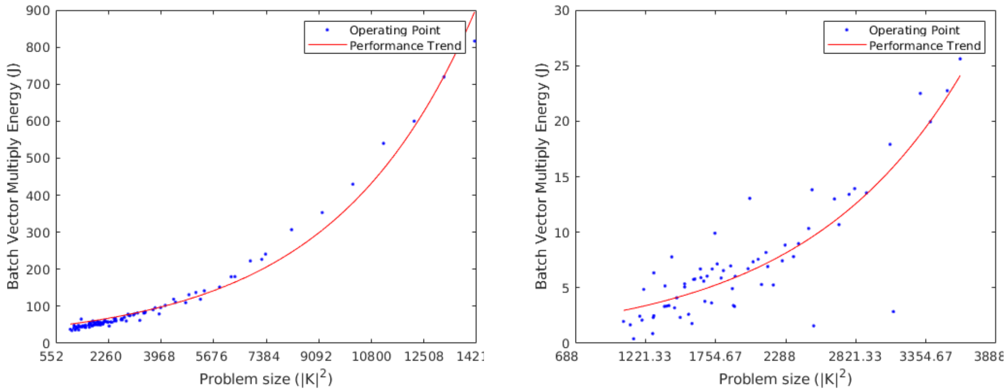
Figure 6.7: Power trace analysis performed by our framework. As the accelerator physical simulation accuracy and size parameters are swept, the framework gathers power traces from the platform under study. Traces are organized into a 2D matrix of increasing liver size and simulation resolution respectively. The power(fig. 6.7b) or energy (fig. 6.7a) drain of the accelerator per second is projected in the third dimension for CAIG frame rates of interest. Each surface in fig. 6.9 shows performance for a specified CAIG simulation frame rate. Next, *platform* battery and thermal models are applied to the frame rate trends to find the optimal contour for each frame rate (indicated by a red intersecting plane). Finally, the contours are re projected into 2D to show optimum platform thermal or battery performance for a given liver procedure as shown in fig. 6.9.

6.5 Results

We wanted to use our method to know what CAIG accuracy and frame rates were possible given a liver procedure and a set of potential CAIG platforms. We specified a 3 hour resection where the CAIG platform temperature should not exceed 45°C. The physical simulation was defined using problem sizes from fig. 6.4 and acceleration benchmarks were created according to the mapping described in section 6.3.1. We then used the method described in 6.3 to evaluate CAIG platforms described in table 6.2. Finally, we performed a conventional performance analysis to judge the relative effectiveness of our method.



(a) K40 GPU Energy Trend (b) DE5 FPGA Energy Trend



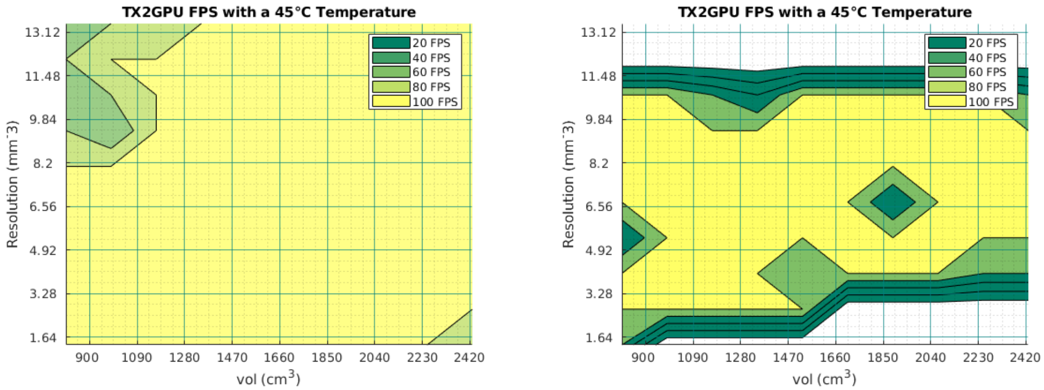
(c) TX2 GPU Energy Trend (d) DE10 FPGA Energy Trend

Figure 6.8: Measured accelerator energy per liver simulation problem iteration with increasing problem sizes for back and head mounted accelerators listed in table 6.2. An iteration is the multiplication of a global stiffness matrix with the simulated degrees of freedom as described in section 6.3.1. The general trend over number of nodes is given in figs. 6.8a to 6.8c.

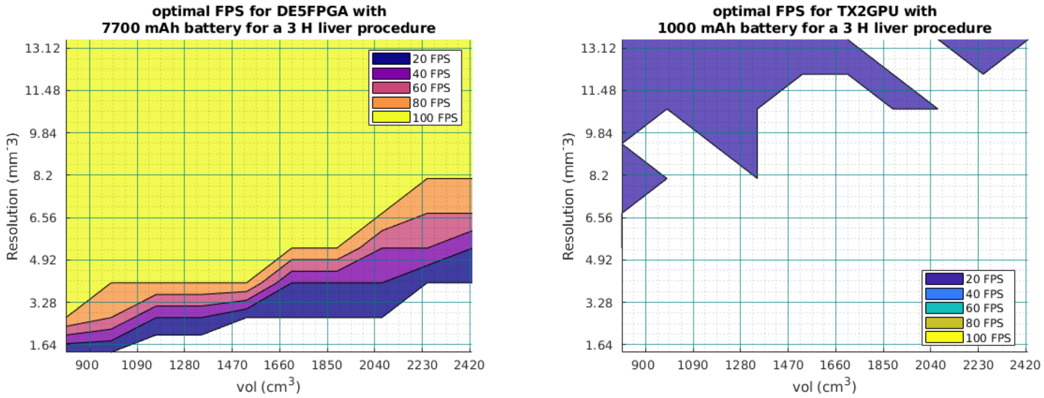
In a conventional analysis, we attempted to find trends in platform power versus problem size as shown in fig. 6.8. In general each Accelerator had a steeply increasing trend with problem size owing to the O(n^3) complexity of the batch vector operation. However simple curve fitting cannot describe performance of accelerators in a clinically relevant way. Performance trends were not simple owing to hardware specifics for each accelerator such as memory transaction scheduling, tiling of the three tuple DOF vectors within accelerator multiply accumulate blocks, and a raft of other factors. On the contrary our method gave helpful colour coded tables on possible frame rates and accuracies for each platform. These results could be reported as a temperature analysis or battery analysis depending on the clinicians priorities.

Platform Temperature Study**:** We assumed an OR was 18°C and that a device hotter than 45°C would be distracting to the clinician. Thermal performance of CAIG platforms was taken from datasheets or could be otherwise estimated [244, 245, 246, 247]. Our analysis found all back mounted devices to be cool enough at all frame rates we considered, however both head mounted devices were estimated to overheat in some cases as shown in Figures 6.9a and 6.9b. These figures are interesting as they show the details of physics simulation acceleration implementation can cause "hot spots" within the design space.

Platform Battery Life Study**:** We assumed head mounted platforms would use smart watch type batteries to be as unobtrusive as possible. Since backpack devices can be larger we assumed laptop type battery capacities. Battery parameters were taken from datasheets [248, 249]. Our analysis found the head mounted FPGA platform unable to last for a 3 hour procedure and that the back mounted GPU platform provided inferior frame rates compared to the back mounted FPGA. This information was combined with our temperature analysis to find the over all best CAIG platform configurations as shown in figs. 6.9c and 6.9d. Figure 6.9c shows the available frame rates for a 3 hour liver procedure with a back mounted device. Figure 6.9d shows frame rates for a head mounted device. Our results were interesting since although FPGAs are considered to be more efficient than GPUs, we found the opposite for the head mounted platforms.



(a) Head mounted GPU Temperature performance (b) Head mounted FPGA temperature performance



(c) Back mounted FPGA battery performance (d) Head mounted GPU battery performance

Figure 6.9: Measured CAIG platform performance with increasing problem sizes for back and head mounted GPU and FPGA devices. Power traces are taken during the multiplication of a global stiffness matrix with the simulated degrees of freedom as described in section 6.3.1. 200 CG iterations are assumed per frame. Simple thermal and battery models are then applied to the measurements. Figures 6.9a and 6.9b are temperature trends for head mounted devices with a 45°C upper bound since these were modeled as overheating in some cases. Figures 6.9c and 6.9d are battery life trends for a back and headset CAIG device respectively. We show the optimal plots for the best overall performing back mounted and head mounted device respectively.

Discussion: Our framework for Efficient Physical Simulation Hardware Acceleration Platforms provided a clinically relevant way of specifying an efficient and unobtrusive CAIG headset from a candidate set of hardware. Unlike conventional analysis, our tools allow simple questions on if a headset will overheat and if the battery will last when surgical procedural details are supplied. Our tool also highlighted the risks of assuming one design will be more efficient than another since our embedded FPGA platform had unexpectedly high static power consumption.

6.6 Conclusions

We introduced a framework for designing efficient head mounted CAIG platforms. Our framework is novel in that it emphasizes the design of clinically useful platforms and that it focuses on physical simulation for accurate placement of guidance cues. The framework comprises of hardware instrumentation methods, a physical simulation acceleration model for multiple accelerator architectures, a CAIG hardware benchmark for multiple accelerators and a specialized power trace analysis tool. Our analysis tool features simple models to convert abstract electrical metrics to practical estimates of battery life and device temperature. We anticipate the framework to have applications in helping surgeons to choose the best compromise between obtrusive CAIG platforms with high performance and unobtrusive CAIG platforms with lower performance when planning CAIG procedures. In future our framework may be used to evaluate other CAIG platforms as the analysis tool, benchmarking codes and a BOM for our hardware monitoring solutions are available under the BSD license [1].

---

[1] Our codes are provided under the BSD license at: https://bitbucket.org/michael_barrow/embeddedarpathfinding/src/master/

Bibliography

[1] Robust nonrigid registration by convex optimization: Complete code. https://drive.google.com/file/d/0B3j4PjSVMIHLYWttRks5N0tmTXM/view?usp=sharing. Accessed: 2017-03-30.

[2] Skand Hurkat, Jungwook Choi, Eriko Nurvitadhi, José F Martínez, and Rob A Rutenbar. Fast hierarchical implementation of sequential tree-reweighted belief propagation for probabilistic inference. In *Field Programmable Logic and Applications (FPL), 2015 25th International Conference on*, pages 1–8. IEEE, 2015.

[3] Chia-Kai Liang, Chao-Chung Cheng, Yen-Chieh Lai, Liang-Gee Chen, and Homer H Chen. Hardware-efficient belief propagation. *IEEE Transactions on Circuits and Systems for Video Technology*, 21(5):525–537, 2011.

[4] Jungwook Choi and Rob A Rutenbar. Video-rate stereo matching using markov random field trw-s inference on a hybrid cpu+ fpga computing platform. *IEEE Transactions on Circuits and Systems for Video Technology*, 26(2):385–398, 2016.

[5] Wenlai Zhao, Haohuan Fu, Guangwen Yang, and Wayne Luk. Patra: Parallel tree-reweighted message passing architecture. In *Field Programmable Logic and Applications (FPL), 2014 24th International Conference on*, pages 1–6. IEEE, 2014.

[6] Vladimir Kolmogorov. Convergent tree-reweighted message passing for energy minimization. *IEEE transactions on pattern analysis and machine intelligence*, 28(10):1568–1583, 2006.

[7] Chengli Song, Afshin Alijani, Tim Frank, GB Hanna, and Alfred Cuschieri. Mechanical properties of the human abdominal wall measured in vivo during insufflation for laparoscopic surgery. *Surgical Endoscopy And Other Interventional Techniques*, 20(6):987–990, 2006.

[8] Eddie K Abdalla, Alban Denys, Patrick Chevalier, Rabih A Nemr, and Jean-Nicolas Vauthey. Total and segmental liver volume variations: implications for liver surgery. *Surgery*, 135(4):404–410, 2004.

[9] World Health Organization. Projections of mortality and causes of death, 2016 to 2060.

[10] International Agency for Research on Cancer, World Health Organization. Cancer today.

[11] Rebecca L Siegel, Kimberly D Miller, and Ahmedin Jemal. Cancer statistics, 2019. *CA: a cancer journal for clinicians*, 69(1):7–34, 2019.

[12] Ahmedin Jemal, Elizabeth M Ward, Christopher J Johnson, Kathleen A Cronin, Jiemin Ma, A Blythe Ryerson, Angela Mariotto, Andrew J Lake, Reda Wilson, Recinda L Sherman, et al. Annual report to the nation on the status of cancer, 1975–2014, featuring survival. *JNCI: Journal of the National Cancer Institute*, 109(9):djx030, 2017.

[13] Lola Rahib, Benjamin D Smith, Rhonda Aizenberg, Allison B Rosenzweig, Julie M Fleshman, and Lynn M Matrisian. Projecting cancer incidence and deaths to 2030: the unexpected burden of thyroid, liver, and pancreas cancers in the united states. *Cancer research*, 74(11):2913–2921, 2014.

[14] Richard N Berri and Eddie K Abdalla. Curable metastatic colorectal cancer: recommended paradigms. *Current oncology reports*, 11(3):200–208, 2009.

[15] Steven R Alberts and Graeme J Poston. Treatment advances in liver-limited metastatic colorectal cancer. *Clinical colorectal cancer*, 10(4):258–265, 2011.

[16] Scott Kopetz and Jean-Nicolas Vauthey. Perioperative chemotherapy for resectable hepatic metastases. *The Lancet*, 371(9617):963–965, 2008.

[17] Eddie K Abdalla, Jean-Nicolas Vauthey, Lee M Ellis, Vickie Ellis, Raphael Pollock, Kristine R Broglio, Kenneth Hess, and Steven A Curley. Recurrence and outcomes following hepatic resection, radiofrequency ablation, and combined resection/ablation for colorectal liver metastases. *Annals of surgery*, 239(6):818, 2004.

[18] S Manfredi, C Lepage, C Hatem, O Coatmeur, J Faivre, and AM Bouveir. Epidemiology and management of colorectal liver metastases. *Ann. Surg*, 244:254–259, 2006.

[19] MJ Schindl, DN Redhead, KCH Fearon, OJ Garden, and SJ Wigmore. The value of residual liver volume as a predictor of hepatic dysfunction and infection after major liver resection. *Gut*, 54(2):289–296, 2005.

[20] Eben L Rosenthal, Jason M Warram, Kirby I Bland, and Kurt R Zinn. The status of contemporary image-guided modalities in oncologic surgery. *Annals of surgery*, 261(1):46, 2015.

[21] Hans-Peter Meinzer, Matthias Thorn, and Carlos E Cárdenas. Computerized planning of liver surgery—an overview. *Computers & Graphics*, 26(4):569–576, 2002.

[22] D Mutter, L Soler, and J Marescaux. Recent advances in liver imaging. *Expert review of gastroenterology & hepatology*, 4(5):613–621, 2010.

[23] H Bourquain, A Schenk, F Link, B Preim, G Prause, and H-O Peitgen. Hepavision2—a software assistant for preoperative planning in living-related liver transplantation and oncologic liver surgery. In *CARS 2002 Computer Assisted Radiology and Surgery*, pages 341–346. Springer, 2002.

[24] Bernhard Reitinger, Alexander Bornik, Reinhard Beichel, Georg Werkgartner, and Erich Sorantin. Tools for augmented-reality-based liver resection planning. In *Medical Imaging 2004: Visualization, Image-Guided Procedures, and Display*, volume 5367, pages 88–99. International Society for Optics and Photonics, 2004.

[25] Hauke Lang, Arnold Radtke, Milo Hindennach, Tobias Schroeder, Nils R Frühauf, Massimo Malagó, Holger Bourquain, Heinz-Otto Peitgen, Karl J Oldhafer, and Christoph E Broelsch. Impact of virtual tumor resection and computer-assisted risk analysis on operation planning and intraoperative strategy in major hepatic resection. *Archives of surgery*, 140(7):629–638, 2005.

[26] Luc Soler, Hervé Delingette, Grégoire Malandain, Johan Montagnat, Nicholas Ayache, Christophe Koehl, Olivier Dourthe, Benoit Malassagne, Michelle Smith, Didier Mutter, et al. Fully automatic anatomical, pathological, and functional segmentation from ct scans for hepatic surgery. *Computer Aided Surgery*, 6(3):131–142, 2001.

[27] V Sojar, D Stanisavljević, M Hribernik, M Glušič, D Kreuh, U Velkavrh, and T Fius. Liver surgery training and planning in 3d virtual space. In *International Congress Series*, volume 1268, pages 390–394. Elsevier, 2004.

[28] Pablo Lamata, Félix Lamata, Valentin Sojar, Piotr Makowski, Laurent Massoptier, Sergio Casciaro, Wajid Ali, Thomas Stüdeli, Jérôme Declerck, Ole Jackov Elle, et al. Use of the resection map system as guidance during hepatectomy. *Surgical endoscopy*, 24(9):2327–2337, 2010.

[29] Itaru Endo, Hiroshi Shimada, Mitsutaka Sugita, Yoshiro Fujii, Daisuke Morioka, Kazuhisa Takeda, Sadatoshi Sugae, Kuniya Tanaka, Shinji Togo, Holger Bourquain, et al. Role of three-dimensional imaging in operative planning for hilar cholangiocarcinoma. *Surgery*, 142(5):666–675, 2007.

[30] Andrea Schenk, Dieter Haemmerich, and Tobias Preusser. Planning of image-guided interventions in the liver. *IEEE pulse*, 2(5):48–55, 2011.

[31] Renate Hammerstingl, Alexander Huppertz, Josy Breuer, Thomas Balzer, Anthony Blakeborough, Rick Carter, Lluis Castells Fusté, Gertraud Heinz-Peer, Werner Judmaier, Michael Laniado, et al. Diagnostic efficacy of gadoxetic acid (primovist)-enhanced mri and spiral ct for a therapeutic strategy: comparison with intraoperative and histopathologic findings in focal liver lesions. *European radiology*, 18(3):457, 2008.

[32] Vanessa Kulemann, Wolfgang Schima, Dietmar Tamandl, Klaus Kaczirek, Thomas Gruenberger, Friedrich Wrba, Michael Weber, and Ahmed Ba-Ssalamah. Preoperative detection

of colorectal liver metastases in fatty liver: Mdct or mri? *European journal of radiology*, 79(2):e1–e6, 2011.

[33] Konstantin Holzapfel, Carolin Reiser-Erkan, Alexander A Fingerle, Merd Erkan, Matthias J Eiber, Ernst J Rummeny, Helmut Friess, Jörg Kleeff, and Jochen Gaa. Comparison of diffusion-weighted mr imaging and multidetector-row ct in the detection of liver metastases in patients operated for pancreatic cancer. *Abdominal imaging*, 36(2):179–184, 2011.

[34] Tejas Parikh, Stephen J Drew, Vivian S Lee, Samson Wong, Elizabeth M Hecht, James S Babb, and Bachir Taouli. Focal liver lesion detection and characterization with diffusion-weighted mr imaging: comparison with standard breath-hold t2-weighted imaging. *Radiology*, 246(3):812–822, 2008.

[35] Kathryn J Fowler, Harmeet Kaur, Brooks D Cash, Barry W Feig, Kenneth L Gage, Evelyn M Garcia, Amy K Hara, Joseph M Herman, David H Kim, Drew L Lambert, et al. Acr appropriateness criteria® pretreatment staging of colorectal cancer. *Journal of the American College of Radiology*, 14(5):S234–S244, 2017.

[36] Yingzhen N Zhang, Kathryn J Fowler, Gavin Hamilton, Jennifer Y Cui, Ethan Z Sy, Michelle Balanay, Jonathan C Hooker, Nikolaus Szeverenyi, and Claude B Sirlin. Liver fat imaging—a clinical overview of ultrasound, ct, and mr imaging. *The British journal of radiology*, 91(1089):20170959, 2018.

[37] Yingzhen N Zhang, Kathryn J Fowler, Arinc Ozturk, Chetan K Potu, Ashley L Louie, Vivian Montes, Walter C Henderson, Kang Wang, Michael P Andre, Anthony E Samir, et al. Liver fibrosis imaging: A clinical review of ultrasound and magnetic resonance elastography. *Journal of Magnetic Resonance Imaging*, 2019.

[38] Tetsuo Ikeda, Yohei Mano, Kazutoyo Morita, Naotaka Hashimoto, Hirohito Kayashima, Atsuro Masuda, Toru Ikegami, Tomoharu Yoshizumi, Ken Shirabe, and Yoshihiko Maehara. Pure laparoscopic hepatectomy in semiprone position for right hepatic major resection. *Journal of Hepato-Biliary-Pancreatic Sciences*, 20(2):145–150, 2013.

[39] Go Wakabayashi, Daniel Cherqui, David A Geller, Joseph F Buell, Hironori Kaneko, Ho Seong Han, Horacio Asbun, Nicholas O'Rourke, Minoru Tanabe, Alan J Koffron, et al. Recommendations for laparoscopic liver resection: a report from the second international consensus conference held in morioka. *Annals of surgery*, 261(4):619–629, 2015.

[40] Igor Peterlik, Hadrien Courtecuisse, Robert Rohling, Purang Abolmaesumi, Christopher Nguan, Stéphane Cotin, and Septimiu Salcudean. Fast elastic registration of soft tissues under large deformations. *Medical image analysis*, 45:24–40, 2018.

[41] William R Jarnagin, Amber L Simpson, and Michael I Miga. Toward integrated image guided liver surgery. In *Medical Imaging 2017: Image-Guided Procedures, Robotic Interventions, and Modeling*, volume 10135, page 101350R. International Society for Optics and Photonics, 2017.

[42] Oleg Heizmann, Stephan Zidowitz, Holger Bourquain, Silke Potthast, Heinz-Otto Peitgen, Daniel Oertli, and Christoph Kettelhack. Assessment of intraoperative liver deformation during hepatic resection: prospective clinical study. *World journal of surgery*, 34(8):1887–1893, 2010.

[43] A Jalote-Parmar, PMT Pattynama, H De Ridder, RHM Goossens, A Freudenthal, and E Samset. Surgical workflow analysis: identifying user requirements for surgical information systems. In *Meeting Diversity in Ergonomics*, pages 229–241, 2007.

[44] Sylvain Bernhardt, Stéphane A Nicolau, Luc Soler, and Christophe Doignon. The status of augmented reality in laparoscopic surgery as of 2016. *Medical image analysis*, 37:66–90, 2017.

[45] Guang-Zhong Yang, David J Hawkes, Daniel Rueckert, Alison Noble, and Chris Taylor. *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2009: 12th International Conference, London, UK, September 20-24, 2009, Proceedings*, volume 5761. Springer, 2009.

[46] Sven Haase, Sebastian Bauer, Jakob Wasza, Thomas Kilgus, Lena Maier-Hein, Armin Schneider, Michael Kranzfelder, Hubertus Feußner, and Joachim Hornegger. 3-d operation situs reconstruction with time-of-flight satellite cameras using photogeometric data fusion. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 356–363. Springer, 2013.

[47] Benjamin J Dixon, Michael J Daly, Harley HL Chan, Allan Vescan, Ian J Witterick, and Jonathan C Irish. Inattentional blindness increased with augmented reality surgical navigation. *American journal of rhinology & allergy*, 28(5):433–437, 2014.

[48] Maria Luz, Gero Strauss, and Dietrich Manzey. Impact of image-guided surgery on surgeons' performance: a literature review. *International Journal of Human Factors and Ergonomics*, 4(3-4):229–263, 2016.

[49] Wolfgang Schima, Christiane Kulinna, Herbert Langenberger, and Ahmed Ba-Ssalamah. Liver metastases of colorectal cancer: Us, ct or mr?: Wednesday 5 october 2005, 14: 00–16: 00. *Cancer Imaging*, 5(Spec No A):S149, 2005.

[50] C Karlo, CS Reiner, P Stolzmann, S Breitenstein, B Marincek, D Weishaupt, and T Frauenfelder. Ct-and mri-based volumetry of resected liver specimen: comparison to intraoperative volume and weight measurements and calculation of conversion factors. *European journal of radiology*, 75(1):e107–e111, 2010.

[51] B Vinoth Kumar, S Sabareeswaran, and G Madumitha. A decennary survey on artificial intelligence methods for image segmentation. In *Advanced Engineering Optimization Through Intelligent Techniques*, pages 291–311. Springer, 2020.

[52] Patrick Ferdinand Christ, Florian Ettlinger, Felix Grün, Mohamed Ezzeldin A Elshaera, Jana Lipkova, Sebastian Schlecht, Freba Ahmaddy, Sunil Tatavarty, Marc Bickel, Patrick Bilic, et al. Automatic liver and tumor segmentation of ct and mri volumes using cascaded fully convolutional neural networks. *arXiv preprint arXiv:1702.05970*, 2017.

[53] Xiao Han. Automatic liver lesion segmentation using a deep convolutional neural network method. *arXiv preprint arXiv:1704.07239*, 2017.

[54] Ye-zhan Zeng, Sheng-hui Liao, Ping Tang, Yu-qian Zhao, Miao Liao, Yan Chen, and Yi-xiong Liang. Automatic liver vessel segmentation using 3d region growing and hybrid active contour model. *Computers in biology and medicine*, 97:63–73, 2018.

[55] Dong Yang, Daguang Xu, S Kevin Zhou, Bogdan Georgescu, Mingqing Chen, Sasa Grbic, Dimitris Metaxas, and Dorin Comaniciu. Automatic liver segmentation using an adversarial image-to-image network. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 507–515. Springer, 2017.

[56] Hieu Trung Huynh, Ibrahim Karademir, Aytekin Oto, and Kenji Suzuki. Computerized liver volumetry on mri by using 3d geodesic active contour segmentation. *American journal of roentgenology*, 202(1):152–159, 2014.

[57] Xuehu Wang, Jian Yang, Danni Ai, Yongchang Zheng, Songyuan Tang, and Yongtian Wang. Adaptive mesh expansion model (amem) for liver segmentation from ct image. *PloS one*, 10(3), 2015.

[58] Jinao Zhang, Yongmin Zhong, and Chengfan Gu. Deformable models for surgical simulation: a survey. *IEEE reviews in biomedical engineering*, 11:143–164, 2017.

[59] Alan Liu, Frank Tendick, Kevin Cleary, and Christoph Kaufmann. A survey of surgical simulation: applications, technology, and education. *Presence: Teleoperators & virtual environments*, 12(6):599–614, 2003.

[60] Atul Thakur, Ashis Gopal Banerjee, and Satyandra K Gupta. A survey of cad model simplification techniques for physics-based simulation applications. *Computer-Aided Design*, 41(2):65–80, 2009.

[61] Qiang Du and Desheng Wang. Recent progress in robust and quality delaunay mesh generation. *Journal of Computational and Applied Mathematics*, 195(1-2):8–23, 2006.

[62] Kenji Shimada and David C Gossard. Bubble mesh: automated triangular meshing of non-manifold geometry by sphere packing. In *Proceedings of the third ACM symposium on Solid modeling and applications*, pages 409–419, 1995.

[63] Daniel SH Lo. *Finite element mesh generation.* CRC Press, 2014.

[64] Jonathan D Spratt, Lonie R Salkowski, Jamie Weir, and Peter H Abrahams. *Imaging Atlas of Human Anatomy E-Book.* Elsevier Health Sciences, 2010.

[65] Ichiro Sakuma, Yosuke Nishimura, Chee Kong Chui, Etsuko Kobayashi, Hiroshi Inada, Xian Chen, and Toshiaki Hisada. In vitro measurement of mechanical properties of liver tissue under compression and elongation using a new test piece holding method with surgical glue. In *International Symposium on Surgery Simulation and Soft Tissue Modeling*, pages 284–292. Springer, 2003.

[66] Alessandro Nava, Edoardo Mazza, M Furrer, Peter Villiger, and WH Reinhart. In vivo mechanical characterization of human liver. *Medical image analysis*, 12(2):203–216, 2008.

[67] Sagar Umale, Caroline Deck, Nicolas Bourdet, Parag Dhumane, Luc Soler, Jacques Marescaux, and Remy Willinger. Experimental mechanical characterization of abdominal organs: liver, kidney & spleen. *Journal of the mechanical behavior of biomedical materials*, 17:22–33, 2013.

[68] Sebastian Mueller and Laurent Sandrin. Liver stiffness: a novel parameter for the diagnosis of liver disease. *Hepatic medicine: evidence and research*, 2:49, 2010.

[69] Nazim Haouchine, Stephane Cotin, Igor Peterlik, Jeremie Dequidt, Mario Sanz Lopez, Erwan Kerrien, and Marie-Odile Berger. Impact of soft tissue heterogeneity on augmented reality for liver surgery. *IEEE Transactions on Visualization & Computer Graphics*, (1):1–1, 2015.

[70] Rosalie Plantefeve, Igor Peterlik, Nazim Haouchine, and Stéphane Cotin. Patient-specific biomechanical modeling for guidance during minimally-invasive hepatic surgery. *Annals of biomedical engineering*, 44(1):139–153, 2016.

[71] Nazim Haouchine, Jeremie Dequidt, Igor Peterlik, Erwan Kerrien, Marie-Odile Berger, and Stéphane Cotin. Image-guided simulation of heterogeneous tissue deformation for augmented reality during hepatic surgery. In *2013 IEEE international symposium on mixed and augmented reality (ISMAR)*, pages 199–208. IEEE, 2013.

[72] Ramya Rao Basava, Jiun-Shyan Chen, Yantao Zhang, Shantanu Sinha, Usha Sinha, John Hodgson, Robert Csapo, and Vadim Malis. Pixel based meshfree modeling of skeletal muscles. In *International Symposium Computational Modeling of Objects Represented in Images*, pages 316–327. Springer, 2014.

[73] Jiun-Shyan Chen, Ramya Rao Basava, Yantao Zhang, Robert Csapo, Vadim Malis, Usha Sinha, John Hodgson, and Shantanu Sinha. Pixel-based meshfree modelling of skeletal muscles. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, 4(2):73–85, 2016.

[74] Nazim Haouchine, Jeremie Dequidt, Igor Peterlik, Erwan Kerrien, Marie-Odile Berger, and Stephane Cotin. Towards an accurate tracking of liver tumors for augmented reality in robotic assisted surgery. In *2014 IEEE International Conference on Robotics and Automation (ICRA)*, pages 4121–4126. IEEE, 2014.

[75] Vincent Wai-Sun Wong, Julien Vergniol, Grace Lai-Hung Wong, Juliette Foucher, Henry Lik-Yuen Chan, Brigitte Le Bail, Paul Cheung-Lung Choi, Mathurin Kowo, Anthony Wing-Hung Chan, Wassil Merrouche, et al. Diagnosis of fibrosis and cirrhosis using liver stiffness measurement in nonalcoholic fatty liver disease. *Hepatology*, 51(2):454–462, 2010.

[76] Donald B Gennery. Stereo-camera calibration. In *Proceedings ARPA IUS Workshop*, pages 101–107, 1979.

[77] Juyang Weng, Paul Cohen, and Marc Herniou. Camera calibration with distortion models and accuracy evaluation. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, (10):965–980, 1992.

[78] Bo Li, Lionel Heng, Kevin Koser, and Marc Pollefeys. A multiple-camera system calibration toolbox using a feature descriptor-based calibration pattern. In *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1301–1307. IEEE, 2013.

[79] AJ Lacey, N Pinitkarn, and Neil A Thacker. An evaluation of the performance of ransac algorithms for stereo camera calibrarion. In *BMVC*, pages 1–10, 2000.

[80] Anne-Sophie Poulin-Girard, Simon Thibault, and Denis Laurendeau. Influence of camera calibration conditions on the accuracy of 3d reconstruction. *Optics express*, 24(3):2678–2686, 2016.

[81] Ramin Shahidi, Michael R Bax, Calvin R Maurer, Jeremy A Johnson, Eric P Wilkinson, Bai Wang, Jay B West, Martin J Citardi, Kim H Manwaring, and Rasool Khadem. Implementation, calibration and accuracy testing of an image-enhanced endoscopy system. *IEEE Transactions on Medical imaging*, 21(12):1524–1535, 2002.

[82] Tetsuzo Yamaguchi, Masahiko Nakamoto, Yoshinobu Sato, Yoshikazu Nakajima, Kozo Konishi, Makoto Hashizume, Takashi Nishii, Nobuhiko Sugano, Hideki Yoshikawa, Kazuo Yonenobu, et al. Camera model and calibration procedure for oblique-viewing endoscope. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 373–381. Springer, 2003.

[83] Radu Horaud and Fadi Dornaika. Hand-eye calibration. *The international journal of robotics research*, 14(3):195–210, 1995.

[84] Logan W Clements, William C Chapman, Benoit M Dawant, Robert L Galloway Jr, and Michael I Miga. Robust surface registration using salient anatomical features for image-guided liver surgery: algorithm and validation. *Medical physics*, 35(6Part1):2528–2540, 2008.

[85] Thiago Ramos dos Santos, Alexander Seitel, Thomas Kilgus, Stefan Suwelack, Anna-Laura Wekerle, Hannes Kenngott, Stefanie Speidel, Heinz-Peter Schlemmer, Hans-Peter

Meinzer, Tobias Heimann, et al. Pose-independent surface matching for intra-operative soft-tissue marker-less registration. *Medical image analysis*, 18(7):1101–1114, 2014.

[86] D Caleb Rucker, Yifei Wu, Logan W Clements, Janet E Ondrake, Thomas S Pheiffer, Amber L Simpson, William R Jarnagin, and Michael I Miga. A mechanics-based nonrigid registration method for liver surgery using sparse intraoperative data. *IEEE transactions on medical imaging*, 33(1):147–158, 2013.

[87] Stefan Suwelack, Sebastian Röhl, Sebastian Bodenstedt, Daniel Reichard, Rüdiger Dillmann, Thiago dos Santos, Lena Maier-Hein, Martin Wagner, Josephine Wünscher, Hannes Kenngott, et al. Physics-based shape matching for intraoperative image guidance. *Medical physics*, 41(11):111901, 2014.

[88] Rosalie Plantefeve, Nazim Haouchine, Jean-Pierre Radoux, and Stéphane Cotin. Automatic alignment of pre and intraoperative data using anatomical landmarks for augmented laparoscopic liver surgery. In *International Symposium on Biomedical Simulation*, pages 58–66. Springer, 2014.

[89] Maria R Robu, Philip Edwards, João Ramalhinho, Stephen Thompson, Brian Davidson, David Hawkes, Danail Stoyanov, and Matthew J Clarkson. Intelligent viewpoint selection for efficient ct to video registration in laparoscopic liver surgery. *International journal of computer assisted radiology and surgery*, 12(7):1079–1088, 2017.

[90] Yi Song, Johannes Totz, Steve Thompson, Stian Johnsen, Dean Barratt, Crispin Schneider, Kurinchi Gurusamy, Brian Davidson, Sébastien Ourselin, David Hawkes, et al. Locally rigid, vessel-based registration for laparoscopic liver surgery. *International journal of computer assisted radiology and surgery*, 10(12):1951–1961, 2015.

[91] Ben Bellekens, Vincent Spruyt, Rafael Berkvens, and Maarten Weyn. A survey of rigid 3d pointcloud registration algorithms. In *AMBIENT 2014: the Fourth International Conference on Ambient Computing, Applications, Services and Technologies, August 24-28, 2014, Rome, Italy*, pages 8–13, 2014.

[92] Anh Nguyen and Bac Le. 3d point cloud segmentation: A survey. In *2013 6th IEEE conference on robotics, automation and mechatronics (RAM)*, pages 225–230. IEEE, 2013.

[93] François Pomerleau, Francis Colas, Roland Siegwart, et al. A review of point cloud registration algorithms for mobile robotics. *Foundations and Trends® in Robotics*, 4(1):1–104, 2015.

[94] Sandrine Voros, Jean-Alexandre Long, and Philippe Cinquin. Automatic detection of instruments in laparoscopic images: A first step towards high-level command of robotic endoscopic holders. *The International Journal of Robotics Research*, 26(11-12):1173–1190, 2007.

[95] Austin Reiter, Peter K Allen, and Tao Zhao. Feature classification for tracking articulated surgical tools. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 592–600. Springer, 2012.

[96] Max Allan, Sébastien Ourselin, Steve Thompson, David J Hawkes, John Kelly, and Danail Stoyanov. Toward detection and localization of instruments in minimally invasive surgery. *IEEE Transactions on Biomedical Engineering*, 60(4):1050–1058, 2012.

[97] Krittin Pachtrachai, Max Allan, Vijay Pawar, Stephen Hailes, and Danail Stoyanov. Hand-eye calibration for robotic assisted minimally invasive surgery without a calibration object. In *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 2485–2491. IEEE, 2016.

[98] Max Allan, Steve Thompson, Matthew J Clarkson, Sébastien Ourselin, David J Hawkes, John Kelly, and Danail Stoyanov. 2d-3d pose tracking of rigid instruments in minimally invasive surgery. In *International Conference on Information Processing in Computer-assisted Interventions*, pages 1–10. Springer, 2014.

[99] Zhiqiang Zhang, Lin Zhang, and Guang-Zhong Yang. A computationally efficient method for hand–eye calibration. *International journal of computer assisted radiology and surgery*, 12(10):1775–1787, 2017.

[100] Steven M Seitz, Brian Curless, James Diebel, Daniel Scharstein, and Richard Szeliski. A comparison and evaluation of multi-view stereo reconstruction algorithms. In *2006 IEEE computer society conference on computer vision and pattern recognition (CVPR'06)*, volume 1, pages 519–528. IEEE, 2006.

[101] Li-Ming Su, Balazs P Vagvolgyi, Rahul Agarwal, Carol E Reiley, Russell H Taylor, and Gregory D Hager. Augmented reality during robot-assisted laparoscopic partial nephrectomy: toward real-time 3d-ct to stereoscopic video registration. *Urology*, 73(4):896–900, 2009.

[102] Haoyin Zhou and Jayender Jagadeesan. Real-time dense reconstruction of tissue surface from stereo optical video. *IEEE transactions on medical imaging*, 2019.

[103] Lena Maier-Hein, Peter Mountney, Adrien Bartoli, Haytham Elhawary, D Elson, Anja Groch, Andreas Kolb, Marcos Rodrigues, J Sorger, Stefanie Speidel, et al. Optical techniques for 3d surface reconstruction in computer-assisted laparoscopic surgery. *Medical image analysis*, 17(8):974–996, 2013.

[104] Lena Maier-Hein, Anja Groch, Adrien Bartoli, Sebastian Bodenstedt, G Boissonnat, P-L Chang, NT Clancy, Daniel S Elson, Sven Haase, Eric Heim, et al. Comparative validation of single-shot optical techniques for laparoscopic 3-d surface reconstruction. *IEEE transactions on medical imaging*, 33(10):1913–1930, 2014.

[105] Congying Sui, Jiahao Wu, Zerui Wang, Gan Ma, and Yunhui Liu. A real-time 3d laparoscopic imaging system: Design, method and validation. *IEEE Transactions on Biomedical Engineering*, 2020.

[106] Hugh Durrant-Whyte and Tim Bailey. Simultaneous localization and mapping: part i. *IEEE robotics & automation magazine*, 13(2):99–110, 2006.

[107] Tim Bailey and Hugh Durrant-Whyte. Simultaneous localization and mapping (slam): Part ii. *IEEE robotics & automation magazine*, 13(3):108–117, 2006.

[108] Bingxiong Lin, Yu Sun, Xiaoning Qian, Dmitry Goldgof, Richard Gitlin, and Yuncheng You. Video-based 3d reconstruction, laparoscope localization and deformation recovery for abdominal minimally invasive surgery: a survey. *The International Journal of Medical Robotics and Computer Assisted Surgery*, 12(2):158–178, 2016.

[109] Guofang Xiao, Ester Bonmati, Stephen Thompson, Joe Evans, John Hipwell, Daniil Nikitichev, Kurinchi Gurusamy, Sébastien Ourselin, David J Hawkes, Brian Davidson, et al. Electromagnetic tracking in image-guided laparoscopic surgery: Comparison with optical tracking and feasibility study of a combined laparoscope and laparoscopic ultrasound system. *Medical physics*, 45(11):5094–5104, 2018.

[110] Mads Nielsen, Peter Johansen, Andrew D Jackson, Benny Lautrup, and Søren Hauberg. Brownian warps for non-rigid registration. *Journal of Mathematical Imaging and Vision*, 31(2-3):221, 2008.

[111] Michael J Black and Paul Anandan. The robust estimation of multiple motions: Parametric and piecewise-smooth flow fields. *Computer vision and image understanding*, 63(1):75–104, 1996.

[112] Jingwei Song, Jun Wang, Liang Zhao, Shoudong Huang, and Gamini Dissanayake. Dynamic reconstruction of deformable soft-tissue with stereo scope in minimal invasive surgery. *IEEE Robotics and Automation Letters*, 3(1):155–162, 2017.

[113] Micha Pfeiffer, Carina Riediger, Stefan Leger, Jens-Peter Kühn, Danilo Seppelt, Ralf-Thorsten Hoffmann, Jürgen Weitz, and Stefanie Speidel. Non-rigid volume to surface registration using a data-driven biomechanical model. *arXiv preprint arXiv:2005.14695*, 2020.

[114] J Dargahi, M Parameswaran, and S Payandeh. A micromachined piezoelectric tactile sensor for an endoscopic grasper-theory, fabrication and experiments. *Journal of microelectromechanical systems*, 9(3):329–335, 2000.

[115] S Sokhanvar, M Packirisamy, and J Dargahi. A multifunctional pvdf-based tactile sensor for minimally invasive surgery. *Smart materials and structures*, 16(4):989, 2007.

[116] Hongqiang Sang, Jintian Yun, Reza Monfaredi, Emmanuel Wilson, Hadi Fooladi, and Kevin Cleary. External force estimation and implementation in robotically assisted minimally invasive surgery. *The International Journal of Medical Robotics and Computer Assisted Surgery*, 13(2):e1824, 2017.

[117] Yun-Hsuan Su, Kevin Huang, and Blake Hannaford. Real-time vision-based surgical tool segmentation with robot kinematics prior. In *2018 International Symposium on Medical Robotics (ISMR)*, pages 1–6. IEEE, 2018.

[118] Jinao Zhang, Yongmin Zhong, and Chengfan Gu. Deformable models for surgical simulation: A survey. *IEEE reviews in biomedical engineering*, 11:143–164, 2018.

[119] Maren Freutel, Hendrik Schmidt, Lutz Dürselen, Anita Ignatius, and Fabio Galbusera. Finite element modeling of soft tissues: material models, tissue interaction and challenges. *Clinical Biomechanics*, 29(4):363–372, 2014.

[120] Grand Roman Joldes, Adam Wittek, and Karol Miller. Suite of finite element algorithms for accurate computation of soft tissue deformation for surgical simulation. *Medical image analysis*, 13(6):912–919, 2009.

[121] Olivier Comas, Zeike A Taylor, Jérémie Allard, Sébastien Ourselin, Stéphane Cotin, and Josh Passenger. Efficient nonlinear fem for soft tissue modelling and its gpu implementation within the open source framework sofa. In *International Symposium on Biomedical Simulation*, pages 28–39. Springer, 2008.

[122] S De, J Kim, and MA Srinivasan. A meshless numerical technique for physically based real time medical. *Medicine Meets Virtual Reality 2001: Outer Space, Inner Space, Virtual Space*, 81:113, 2001.

[123] Suvranu De, Jung Kim, Yi-Je Lim, and Mandayam A Srinivasan. The point collocation-based method of finite spheres (pcmfs) for real time surgery simulation. *Computers & structures*, 83(17-18):1515–1525, 2005.

[124] Micha Pfeiffer, Carina Riediger, Jürgen Weitz, and Stefanie Speidel. Learning soft tissue behavior of organs for surgical navigation with convolutional neural networks. *International journal of computer assisted radiology and surgery*, 14(7):1147–1155, 2019.

[125] Qizhi He and Jiun-Shyan Chen. A physics-constrained data-driven approach based on locally convex reconstruction for noisy database. *Computer Methods in Applied Mechanics and Engineering*, 363:112791, 2020.

[126] Stéphane Nicolau, Luc Soler, Didier Mutter, and Jacques Marescaux. Augmented reality in laparoscopic surgical oncology. *Surgical oncology*, 20(3):189–201, 2011.

[127] Nicolas C Buchs, Francesco Volonte, François Pugin, Christian Toso, Matteo Fusaglia, Kate Gavaghan, Pietro E Majno, Matthias Peterhans, Stefan Weber, and Philippe Morel.

Augmented environments for the targeting of hepatic lesions during image-guided robotic liver surgery. *Journal of surgical research*, 184(2):825–831, 2013.

[128] Marta Kersten-Oertel, Pierre Jannin, and D Louis Collins. The state of the art of visualization in mixed reality image guided surgery. *Computerized Medical Imaging and Graphics*, 37(2):98–112, 2013.

[129] Marta Kersten-Oertel, Pierre Jannin, and D Louis Collins. Dvv: a taxonomy for mixed reality visualization in image guided surgery. *IEEE transactions on visualization and computer graphics*, 18(2):332–352, 2011.

[130] Benjamin J Dixon, Michael J Daly, Harley Chan, Allan D Vescan, Ian J Witterick, and Jonathan C Irish. Surgeons blinded by enhanced navigation: the effect of augmented reality on attention. *Surgical endoscopy*, 27(2):454–461, 2013.

[131] Rui Tang, Long-Fei Ma, Zhi-Xia Rong, Mo-Dan Li, Jian-Ping Zeng, Xue-Dong Wang, Hong-En Liao, and Jia-Hong Dong. Augmented reality technology for preoperative planning and intraoperative navigation during hepatobiliary surgery: a review of current methods. *Hepatobiliary & Pancreatic Diseases International*, 17(2):101–112, 2018.

[132] Primoz Markelj, Dejan Tomaževič, Bostjan Likar, and Franjo Pernuš. A review of 3d/2d registration methods for image-guided interventions. *Medical image analysis*, 16(3):642–661, 2012.

[133] Rui Liao, Li Zhang, Ying Sun, Shun Miao, and Christophe Chefd'Hotel. A review of recent advances in registration techniques applied to minimally invasive therapy. *IEEE transactions on multimedia*, 15(5):983–1000, 2013.

[134] Junqing Wang, Weichuan Zhang, and Meirong Chen. A survey of 3d image navigation and high precision dynamic registration in minimally invasive surgery. *Procedia computer science*, 131:320–326, 2018.

[135] Dimitrios Ntourakis, Ricardo Memeo, Luc Soler, Jacques Marescaux, Didier Mutter, and Patrick Pessaux. Augmented reality guidance for the resection of missing colorectal liver metastases: an initial experience. *World journal of surgery*, 40(2):419–426, 2016.

[136] Patrick Pessaux, Michele Diana, Luc Soler, Tullio Piardi, Didier Mutter, and Jacques Marescaux. Towards cybernetic surgery: robotic and augmented reality-assisted liver segmentectomy. *Langenbeck's archives of surgery*, 400(3):381–385, 2015.

[137] Suraj D Serai, Andrew T Trout, Alexander Miethke, Eric Diaz, Stavra A Xanthakos, and Jonathan R Dillman. Putting it all together: established and emerging mri techniques for detecting and measuring liver fibrosis. *Pediatric radiology*, 48(9):1256–1272, 2018.

[138] Andrés Domínguez, Maximiliano Noceti, Daniel Fino, and Pablo Ariza. Hepatic elastography and other advanced mri sequences (multiparametric mri). *Rev. Argent. Radiol*, 82(2):64–71, 2018.

[139] Satoshi Ieiri, Munenori Uemura, Kouzou Konishi, Ryota Souzaki, Yoshihiro Nagao, Norifumi Tsutsumi, Tomohiko Akahoshi, Kenoki Ohuchida, Takeshi Ohdaira, Morimasa Tomikawa, et al. Augmented reality navigation system for laparoscopic splenectomy in children based on preoperative ct image using optical tracking device. *Pediatric surgery international*, 28(4):341–346, 2012.

[140] Shinji Onda, Tomoyoshi Okamoto, Masaru Kanehira, Shuichi Fujioka, Naoki Suzuki, Asaki Hattori, and Katsuhiko Yanaga. Short rigid scope and stereo-scope designed specifically for open abdominal navigation surgery: clinical application for hepatobiliary and pancreatic surgery. *Journal of hepato-biliary-pancreatic sciences*, 20(4):448–453, 2013.

[141] Jędrzej Kowalczuk, Avishai Meyer, Jay Carlson, Eric T Psota, Shelby Buettner, Lance C Pérez, Shane M Farritor, and Dmitry Oleynikov. Real-time three-dimensional soft tissue reconstruction for laparoscopic surgery. *Surgical endoscopy*, 26(12):3413–3417, 2012.

[142] Andreas Schoob, Dennis Kundrat, Lukas Kleingrothe, Lüder A Kahrs, Nicolas Andreff, and Tobias Ortmaier. Tissue surface information for intraoperative incision planning and focus adjustment in laser surgery. *International journal of computer assisted radiology and surgery*, 10(2):171–181, 2015.

[143] Archie Hughes-Hallett, Erik K Mayer, Philip Pratt, Alex Mottrie, Ara Darzi, and Justin Vale. The current and future use of imaging in urological robotic surgery: a survey of the european association of robotic urological surgeons. *The International Journal of Medical Robotics and Computer Assisted Surgery*, 11(1):8–14, 2015.

[144] Francesco Volonté, Francois Pugin, Nicolas Christian Buchs, Joël Spaltenstein, Monika Hagen, Osman Ratib, and Philippe Morel. Console-integrated stereoscopic osirix 3d volume-rendered images for da vinci colorectal robotic surgery. *Surgical innovation*, 20(2):158–163, 2013.

[145] J Michael Fitzpatrick. *Detecting failure, assessing success*. CRC Press New York, 2001.

[146] Jaesung Hong, Nozomu Matsumoto, Riichi Ouchida, Shizuo Komune, and Makoto Hashizume. Medical navigation system for otologic surgery based on hybrid registration and virtual intraoperative computed tomography. *IEEE Transactions on Biomedical Engineering*, 56(2):426–432, 2009.

[147] Raj Shekhar, Omkar Dandekar, Venkatesh Bhat, Mathew Philip, Peng Lei, Carlos Godinez, Erica Sutton, Ivan George, Steven Kavic, Reuben Mezrich, et al. Live augmented reality: a new visualization method for laparoscopic surgery using continuous volumetric computed tomography. *Surgical endoscopy*, 24(8):1976–1985, 2010.

[148] Jochen F. Krucker, Gerald L LeCarpentier, J Brian Fowlkes, and Paul L Carson. Rapid elastic image registration for 3-d ultrasound. *IEEE Transactions on Medical Imaging*, 21(11):1384–1394, 2002.

[149] Igor D Grachev, Dmitriy Berdichevsky, Scott L Rauch, Stephan Heckers, David N Kennedy, Verne S Caviness, and Nathaniel M Alpert. A method for assessing the accuracy of intersubject registration of the human brain using anatomic landmarks. *NeuroImage*, 9(2):250–268, 1999.

[150] Nuwan D Nanayakkara, Bernard Chiu, Abbas Samani, J David Spence, Jagath Samarabandu, and Aaron Fenster. A "twisting and bending" model-based nonrigid image registration technique for 3-d ultrasound carotid images. *IEEE transactions on medical imaging*, 27(10):1378–1388, 2008.

[151] Yuichi Tamura, Nobuhiko Sugano, Toshihiko Sasama, Yoshinobu Sato, Shinichi Tamura, Kazuo Yonenobu, Hideki Yoshikawa, and Takahiro Ochi. Surface-based registration accuracy of ct-based image-guided spine surgery. *European spine journal*, 14(3):291–297, 2005.

[152] E Ishmael Parsai, Komanduri M Ayyangar, Ralph R Dobelbower, and Jeffry A Siegel. Clinical fusion of three-dimensional images using bremsstrahlung spect and ct. *The Journal of Nuclear Medicine*, 38(2):319, 1997.

[153] Sarah J Nelson, Mark R Day, Paul J Buffone, Lawrence L Wald, Thomas F Budinger, Randall Hawkins, William P Dillon, Stephen Huhn, Michael D Prados, Susan Chang, et al. Alignment of volume mr images and high resolution [18f] fluorodeoxyglucose pet images for the evaluation of patients with brain tumors. *Journal of computer assisted tomography*, 21(2):183–191, 1997.

[154] J Michael Fitzpatrick, Jay B West, and Calvin R Maurer. Predicting error in rigid-body point-based registration. *IEEE transactions on medical imaging*, 17(5):694–702, 1998.

[155] Jay B West, J Michael Fitzpatrick, Steven A Toms, Calvin R Maurer Jr, and Robert J Maciunas. Fiducial point placement and the accuracy of point-based, rigid body registration. *Neurosurgery*, 48(4):810–817, 2001.

[156] Christoph Hoffmann, Sonja Krause, Eva Maria Stoiber, Angela Mohr, Stefan Rieken, Oliver Schramm, Jürgen Debus, Florian Sterzing, Rolf Bendl, and Kristina Giske. Accuracy quantification of a deformable image registration tool applied in a clinical setting. *Journal of applied clinical medical physics*, 15(1):237–245, 2014.

[157] Stephen Thompson, Crispin Schneider, Michele Bosi, Kurinchi Gurusamy, Sébastien Ourselin, Brian Davidson, David Hawkes, and Matthew J Clarkson. In vivo estimation of target registration errors during augmented reality laparoscopic surgery. *International journal of computer assisted radiology and surgery*, 13(6):865–874, 2018.

[158] IC London. Hamlyn centre laparoscopic/endoscopic video datasets. *URL http://hamlyn. doc. ic. ac. uk/vision/.[Accessed 15 Jan 2019]*, 2019.

[159] Richard Castillo, Edward Castillo, Rudy Guerra, Valen E Johnson, Travis McPhail, Amit K Garg, and Thomas Guerrero. A framework for evaluation of deformable image registration spatial accuracy using large landmark point sets. *Physics in Medicine & Biology*, 54(7):1849, 2009.

[160] Benedict J Brown and Szymon Rusinkiewicz. Global non-rigid alignment of 3-d scans. In *ACM Transactions on Graphics (TOG)*, volume 26, page 21. ACM, 2007.

[161] Brian Amberg, Sami Romdhani, and Thomas Vetter. Optimal step nonrigid icp algorithms for surface registration. In *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*, pages 1–8. IEEE, 2007.

[162] Qifeng Chen and Vladlen Koltun. Robust nonrigid registration by convex optimization. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 2039–2047, 2015.

[163] Michael Barrow, Steven M Burns, and Ryan Kastner. A fpga accelerator for real-time 3d non-rigid registration using tree reweighted message passing and dynamic markov random field generation. In *2018 28th International Conference on Field Programmable Logic and Applications (FPL)*, pages 335–3357. IEEE, 2018.

[164] Ping-Lin Chang, Danail Stoyanov, Andrew J Davison, et al. Real-time dense stereo reconstruction using convex optimisation with a cost-volume for image-guided robotic surgery. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 42–49. Springer, 2013.

[165] Nicolas Padoy, Tobias Blum, Seyed-Ahmad Ahmadi, Hubertus Feussner, Marie-Odile Berger, and Nassir Navab. Statistical modeling and recognition of surgical workflow. *Medical image analysis*, 16(3):632–641, 2012.

[166] Loubna Bouarfa, Pieter P Jonker, and Jenny Dankelman. Discovery of high-level tasks in the operating room. *Journal of biomedical informatics*, 44(3):455–462, 2011.

[167] Richard A Newcombe, Dieter Fox, and Steven M Seitz. Dynamicfusion: Reconstruction and tracking of non-rigid scenes in real-time. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 343–352, 2015.

[168] Jörg Stückler and Sven Behnke. Multi-resolution surfel maps for efficient dense 3d modeling and tracking. *Journal of Visual Communication and Image Representation*, 25(1):137–147, 2014.

[169] Mao Ye, Huamin Wang, Nianchen Deng, Xubo Yang, and Ruigang Yang. Real-time human pose and shape estimation for virtual try-on using a single commodity depth camera. *IEEE transactions on visualization and computer graphics*, 20(4):550–559, 2014.

[170] Mingsong Dou, Sameh Khamis, Yury Degtyarev, Philip Davidson, Sean Ryan Fanello, Adarsh Kowdle, Sergio Orts Escolano, Christoph Rhemann, David Kim, Jonathan Taylor, et al. Fusion4d: Real-time performance capture of challenging scenes. *ACM Transactions on Graphics (TOG)*, 35(4):114, 2016.

[171] Rad Madhavan, Tsai Hong, and Elena Messina. Temporal range registration for unmanned ground and aerial vehicles. *Journal of Intelligent and Robotic Systems*, 44(1):47–69, 2005.

[172] Jamie Shotton, Andrew Fitzgibbon, Mat Cook, Toby Sharp, Mark Finocchio, Richard Moore, Alex Kipman, and Andrew Blake. Real-time human pose recognition in parts from single depth images. In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pages 1297–1304. Ieee, 2011.

[173] Cem Keskin, Furkan Kıraç, Yunus Emre Kara, and Lale Akarun. Real time hand pose estimation using depth sensors. In *Consumer depth cameras for computer vision*, pages 119–137. Springer, 2013.

[174] Matheen Siddiqui and Gérard Medioni. Human pose estimation from a single view point, real-time range sensor. In *Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on*, pages 1–8. IEEE, 2010.

[175] Michael Zollhöfer, Matthias Nießner, Shahram Izadi, Christoph Rehmann, Christopher Zach, Matthew Fisher, Chenglei Wu, Andrew Fitzgibbon, Charles Loop, Christian Theobalt, et al. Real-time non-rigid reconstruction using an rgb-d camera. *ACM Transactions on Graphics (TOG)*, 33(4):156, 2014.

[176] Robert W Sumner, Johannes Schmid, and Mark Pauly. Embedded deformation for shape manipulation. In *ACM Transactions on Graphics (TOG)*, volume 26, page 80. ACM, 2007.

[177] Richard Szeliski, Ramin Zabih, Daniel Scharstein, Olga Veksler, Vladimir Kolmogorov, Aseem Agarwala, Marshall Tappen, and Carsten Rother. A comparative study of energy minimization methods for markov random fields with smoothness-based priors. *IEEE transactions on pattern analysis and machine intelligence*, 30(6):1068–1080, 2008.

[178] Chisel3. https://github.com/freechipsproject/chisel3.git. Accessed: 2017-03-30.

[179] Rapid design methods for developing hardware accelerators. https://github.com/intel/rapid-design-methods-for-developming-hardware-accelerators.git. Accessed: 2017-03-30.

[180] Federica Bogo, Javier Romero, Matthew Loper, and Michael J Black. Faust: Dataset and evaluation for 3d mesh registration. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3794–3801, 2014.

[181] Roya Rezvani Habibabadi, Pegah Khoshpouri, Maryam Ghadimi, Mohammadreza Shaghaghi, Sanaz Ameli, Bita Hazhirkarzar, Pallavi Pandey, Mounes Aliyari Ghasabeh,

Ankur Pandey, and Ihab R Kamel. Comparison between roi-based and volumetric measurements in quantifying heterogeneity of liver stiffness using mr elastography. *European Radiology*, 30(3):1609–1615, 2020.

[182] Bruno Marques, Rosalie Plantefeve, Frédérick Roy, Nazim Haouchine, Emmanuel Jeanvoine, Igor Peterlik, and Stéphane Cotin. Framework for augmented reality in minimally invasive laparoscopic surgery. In *E-health Networking, Application & Services (HealthCom), 2015 17th International Conference on*, pages 22–27. IEEE, 2015.

[183] Hassan Masoumi, Alireza Behrad, Mohammad Ali Pourmina, and Alireza Roosta. Automatic liver segmentation in mri images using an iterative watershed algorithm and artificial neural network. *Biomedical signal processing and control*, 7(5):429–437, 2012.

[184] Fernando López-Mir, Valery Naranjo, Jesús Angulo, M Alcañiz, and Luis Luna. Liver segmentation in mri: A fully automatic method based on stochastic partitions. *Computer methods and programs in biomedicine*, 114(1):11–28, 2014.

[185] Evgin Goceri, Mehmet Z Unlu, Cuneyt Guzelis, and Oguz Dicle. An automatic level set based liver segmentation from mri data sets. In *2012 3rd international conference on image processing theory, tools and applications (IPTA)*, pages 192–197. IEEE, 2012.

[186] D Engwirda. Unstructured mesh methods for the navier-stokes equations. *Undergraduate Thesis, School of Engineering, University of Sidney*, 2005.

[187] Darren Engwirda. Locally optimal delaunay-refinement and optimisation-based mesh generation. 2014.

[188] François Faure, Benjamin Gilles, Guillaume Bousquet, and Dinesh K Pai. Sparse meshless models of complex deformable solids. In *ACM transactions on graphics (TOG)*, volume 30, page 73. ACM, 2011.

[189] David Baraff and Andrew Witkin. Large steps in cloth simulation. In *Proceedings of the 25th annual conference on Computer graphics and interactive techniques*, pages 43–54. ACM, 1998.

[190] R Muthupillai, DJ Lomas, PJ Rossman, James F Greenleaf, A Manduca, and RL Ehman. Magnetic resonance elastography by direct visualization of propagating acoustic strain waves. *Science*, 269(5232):1854–1857, 1995.

[191] Sudhakar K Venkatesh, Meng Yin, and Richard L Ehman. Magnetic resonance elastography of liver: technique, analysis, and clinical applications. *Journal of magnetic resonance imaging*, 37(3):544–555, 2013.

[192] Kevin J Glaser, Armando Manduca, and Richard L Ehman. Review of mr elastography applications and recent developments. *Journal of Magnetic Resonance Imaging*, 36(4):757–774, 2012.

[193] Armando Manduca, Travis E Oliphant, MA Dresner, JL Mahowald, Scott A Kruse, E Amromin, Joel P Felmlee, James F Greenleaf, and Richard L Ehman. Magnetic resonance elastography: non-invasive mapping of tissue elasticity. *Medical image analysis*, 5(4):237–254, 2001.

[194] Simon Chatelin, Isabelle Charpentier, Nadege Corbin, Laurence Meylheuc, and Jonathan Vappou. An automatic differentiation-based gradient method for inversion of the shear wave equation in magnetic resonance elastography: specific application in fibrous soft tissues. *Physics in Medicine & Biology*, 61(13):5000, 2016.

[195] Matthew C Murphy, Armando Manduca, Joshua D Trzasko, Kevin J Glaser, John Huston III, and Richard L Ehman. Artificial neural networks for stiffness estimation in magnetic resonance elastography. *Magnetic resonance in medicine*, 80(1):351–360, 2018.

[196] Stavra A Xanthakos, Daniel J Podberesky, Suraj D Serai, Lili Miles, Eileen C King, William F Balistreri, and Rohit Kohli. Use of magnetic resonance elastography to assess hepatic fibrosis in children with chronic liver disease. *The Journal of pediatrics*, 164(1):186–188, 2014.

[197] Yu Shi, Qiyong Guo, Fei Xia, Jiaxing Sun, and Yuying Gao. Short-and midterm repeatability of magnetic resonance elastography in healthy volunteers at 3.0 t. *Magnetic resonance imaging*, 32(6):665–670, 2014.

[198] Jiming Zhang, Claudio Arena, Amol Pednekar, Brenda Lambert, Debra Dees, Vei Vei Lee, and Raja Muthupillai. Short-term repeatability of magnetic resonance elastography at 3.0 t: Effects of motion-encoding gradient direction, slice position, and meal ingestion. *Journal of Magnetic Resonance Imaging*, 43(3):704–712, 2016.

[199] Temel Kaya Yasar, Mathilde Wagner, Octavia Bane, Cecilia Besa, James S Babb, Stephan Kannengiesser, Maggie Fung, Richard L Ehman, and Bachir Taouli. Interplatform reproducibility of liver and spleen stiffness measured with mr elastography. *Journal of Magnetic Resonance Imaging*, 43(5):1064–1072, 2016.

[200] Catherine DG Hines, Thorsten A Bley, Mary J Lindstrom, and Scott B Reeder. Repeatability of magnetic resonance elastography for quantification of hepatic stiffness. *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 31(3):725–731, 2010.

[201] Sudhakar K Venkatesh, Gang Wang, Lynette LS Teo, and Bertrand WL Ang. Magnetic resonance elastography of liver in healthy asians: normal liver stiffness quantification and reproducibility assessment. *Journal of Magnetic Resonance Imaging*, 39(1):1–8, 2014.

[202] Zeike A Taylor, Mario Cheng, and Sébastien Ourselin. Real-time nonlinear finite element analysis for surgical simulation using graphics processing units. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 701–708. Springer, 2007.

[203] Everton Hermann, Bruno Raffin, François Faure, Thierry Gautier, and Jérémie Allard. Multi-gpu and multi-cpu parallelization for interactive physics simulations. In *European Conference on Parallel Processing*, pages 235–246. Springer, 2010.

[204] Stian F Johnsen, Zeike A Taylor, Matthew J Clarkson, John Hipwell, Marc Modat, Bjoern Eiben, Lianghao Han, Yipeng Hu, Thomy Mertzanidou, David J Hawkes, et al. Niftysim: A gpu-based nonlinear finite element package for simulation of soft tissue biomechanics. *International journal of computer assisted radiology and surgery*, 10(7):1077–1095, 2015.

[205] Shiyu Jia, Weizhong Zhang, Xiaokang Yu, and Zhenkuan Pan. Cpu–gpu parallel framework for real-time interactive cutting of adaptive octree-based deformable objects. In *Computer Graphics Forum*, volume 37, pages 45–59. Wiley Online Library, 2018.

[206] Christoph J Paulus, Nazim Haouchine, David Cazier, and Stéphane Cotin. Surgical augmented reality with topological changes. In *International conference on medical image computing and computer-assisted intervention*, pages 413–420. Springer, 2015.

[207] Christoph J Paulus, Nazim Haouchine, David Cazier, and Stephane Cotin. Augmented reality during cutting and tearing of deformable objects. In *Mixed and Augmented Reality (ISMAR), 2015 IEEE International Symposium on*, pages 54–59. IEEE, 2015.

[208] Jun Wu, Rüdiger Westermann, and Christian Dick. Physically-based simulation of cuts in deformable bodies: A survey. In *Eurographics (State of the Art Reports)*, pages 1–19. Citeseer, 2014.

[209] Kevin Montgomery, Cynthia Bruyns, Joel Brown, Stephen Sorkin, Frederic Mazzella, Guillaume Thonier, Arnaud Tellier, Benjamin Lerman, and Anil Menon. Spring: A general framework for collaborative, real-time surgical simulation. *Studies in health technology and informatics*, pages 296–303, 2002.

[210] Jérémie Allard, Stéphane Cotin, François Faure, Pierre-Jean Bensoussan, François Poyer, Christian Duriez, Hervé Delingette, and Laurent Grisoni. Sofa-an open source framework for medical simulation. In *MMVR 15-Medicine Meets Virtual Reality*, volume 125, pages 13–18. IOP Press, 2007.

[211] Mafalda Camara, Erik Mayer, Ara Darzi, and Philip Pratt. Soft tissue deformation for surgical simulation: a position-based dynamics approach. *International journal of computer assisted radiology and surgery*, 11(6):919–928, 2016.

[212] Dong Ho Lee, Jeong Min Lee, Joon Koo Han, and Byung Ihn Choi. Mr elastography of healthy liver parenchyma: Normal value and reliability of the liver stiffness value measurement. *Journal of Magnetic Resonance Imaging*, 38(5):1215–1223, 2013.

[213] François Faure, Christian Duriez, Hervé Delingette, Jérémie Allard, Benjamin Gilles, Stéphanie Marchesseau, Hugo Talbot, Hadrien Courtecuisse, Guillaume Bousquet, Igor Peterlik, et al. Sofa: A multi-model framework for interactive physical simulation. In *Soft*

Tissue Biomechanical Modeling for Computer Assisted Surgery, pages 283–321. Springer, 2012.

[214] Christian Hansen, Jan Wieferich, Felix Ritter, Christian Rieder, and Heinz-Otto Peitgen. Illustrative visualization of 3d planning models for augmented reality in liver surgery. *International journal of computer assisted radiology and surgery*, 5(2):133–141, 2010.

[215] Gang Chen, Xue-cheng Li, Guo-qing Wu, Yi Wang, Bin Fang, Xiao-feng Xiong, Ri-gao Yang, Li-wen Tan, Shao-xiang Zhang, and Jia-hong Dong. The use of virtual reality for the functional simulation of hepatic tumors (case control study). *International Journal of Surgery*, 8(1):72–78, 2010.

[216] Kai Siegfried Lehmann, Bernd Benedikt Frericks, Christoph Holmer, Andrea Schenk, Andreas Weihusen, Verena Knappe, Urte Zurbuchen, Heinz Otto Peitgen, Heinz Johannes Buhr, and Jörg Peter Ritz. In vivo validation of a therapy planning system for laser-induced thermotherapy (litt) of liver malignancies. *International journal of colorectal disease*, 26(6):799, 2011.

[217] Shinichi Saito, Junichi Yamanaka, Koui Miura, Norio Nakao, Tomohiro Nagao, Takaaki Sugimoto, Tadamichi Hirano, Nobukazu Kuroda, Yuji Iimuro, and Jiro Fujimoto. A novel 3d hepatectomy simulation based on liver circulation: application to liver resection and transplantation. *Hepatology*, 41(6):1297–1304, 2005.

[218] Stephen J Wigmore, Doris N Redhead, Xue J Yan, John Casey, Krishnakumar Madhavan, Cornelis HC Dejong, Elspeth J Currie, and O James Garden. Virtual hepatic resection using three-dimensional reconstruction of helical computed tomography angioportograms. *Annals of surgery*, 233(2):221, 2001.

[219] Luc Soler and Jacques Marescaux. Patient-specific surgical simulation. *World journal of surgery*, 32(2):208–212, 2008.

[220] Luc Soler, Stephane Nicolau, Patrick Pessaux, Didier Mutter, and Jacques Marescaux. Real-time 3d image reconstruction guidance in liver resection surgery. *Hepatobiliary surgery and nutrition*, 3(2):73, 2014.

[221] Julie Hallet, Brice Gayet, Allan Tsung, Go Wakabayashi, Patrick Pessaux, and 2nd International Consensus Conference on Laparoscopic Liver Resection Group. Systematic review of the use of pre-operative simulation and navigation for hepatectomy: current status and future perspectives. *Journal of Hepato-Biliary-Pancreatic Sciences*, 22(5):353–362, 2015.

[222] Qizhi He and Jiun-Shyan Chen. A physics-constrained data-driven approach based on locally convex reconstruction for noisy database. *arXiv preprint arXiv:1907.12651*, 2019.

[223] Zhengyou Zhang. Iterative point matching for registration of free-form curves and surfaces. *International journal of computer vision*, 13(2):119–152, 1994.

[224] Szymon Rusinkiewicz and Marc Levoy. Efficient variants of the icp algorithm. In *3dim*, volume 1, pages 145–152, 2001.

[225] Andrew J Davison, Ian D Reid, Nicholas D Molton, and Olivier Stasse. Monoslam: Real-time single camera slam. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, (6):1052–1067, 2007.

[226] Richard A Newcombe, Shahram Izadi, Otmar Hilliges, David Molyneaux, David Kim, Andrew J Davison, Pushmeet Kohi, Jamie Shotton, Steve Hodges, and Andrew Fitzgibbon. Kinectfusion: Real-time dense surface mapping and tracking. In *Mixed and augmented reality (ISMAR), 2011 10th IEEE international symposium on*, pages 127–136. IEEE, 2011.

[227] Lutz Nolte, Lucia Zamorano, E Arm, H Visarius, Zaowei Jiang, U Berlerman, and O Schwarzenbach. Image-guided computer-assisted spine surgery: a pilot study on pedicle screw fixation. *Stereotactic and functional neurosurgery*, 66(1-3):108–117, 1996.

[228] Maxwell Lewis Neal and Roy Kerckhoffs. Current progress in patient-specific modeling. *Briefings in bioinformatics*, 11(1):111–126, 2009.

[229] Hadrien Courtecuisse, Stéphane Cotin, Jérémie Allard, and Luc Soler. Gpu-based interactive simulation of liver resection. In *ACM SIGGRAPH 2011 Computer Animation Festival*, pages 98–98. ACM, 2011.

[230] Jérémie Allard, Hadrien Courtecuisse, and François Faure. Implicit fem solver on gpu for interactive deformation simulation. In *GPU computing gems Jade Edition*, pages 281–294. Elsevier, 2012.

[231] Alexander Bornik, Reinhard Beichel, Bernhard Reitinger, Georg Gotschuli, Erich Sorantin, Franz W. Leberl, and Milan Sonka. Computer-aided liver surgery planning: an augmented reality approach. *Medical Imaging 2003*, 2003.

[232] Youngjun Kim, Hannah Kim, and Yong Oock Kim. Virtual reality and augmented reality in plastic surgery: A review. *Archives of Plastic Surgery*, pages 179–187, 2017.

[233] Xiaojun Chen, Lu Xu, Yiping Wang, Huixiang Wang, Fang Wang, Xiangsen Zeng, Qiugen Wang, and Jan Egger. Development of a surgical navigation system based on augmented reality using an optical see-through head-mounted display. *Journal of Biomedical Informatics*, 55:124–131, 2015.

[234] Steven Feiner, Blair MacIntyre, Tobias Höllerer, and Anthony Webster. A touring machine: Prototyping 3d mobile augmented reality systems for exploring the urban environment. *Personal Technologies*, 1:208–217, 1997.

[235] Tobias H. Höllerer and Steven K. Feiner. *Telegeoinformatics: Location-Based Computing and Services: Chapter 9 Mobile Augmented Reality*. Taylor & Francis Books Ltd., 2004.

[236] M. Claudia Leue, Timothy Jung, and Dario tom Dieck. Google glass augmented reality: Generic learning outcomes for art galleries. *Information and Communication Technologies in Tourism*, pages 463–476, 2015.

[237] Janne Paavilainen, Hannu Korhonen, Kati Alha, Jaako Stenros, Elina Koskinen, and Frans Mayra. The pokémon go experience: A location-based augmented reality mobile game goes mainstream. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pages 2493–2498, 2017.

[238] Mohamed Abouzahir, Abdelhafid Elouardi, Rachid Latif, Samir Bouaziz, and Abdelouahed Tajer. Embedding slam algorithms: Has it come of age? *Robotics and Autonomous Systems*, 100:14–26, 2018.

[239] Konstantinos Boikos and Christos-Savvas Bouganis. Semi-dense slam on an fpga soc. In *2016 26th International Conference on Field Programmable Logic and Applications (FPL)*, pages 1–4. IEEE, 2016.

[240] John Dolbow and Ted Belytschko. An introduction to programming the meshless element f reegalerkin method. *Archives of computational methods in engineering*, 5(3):207–241, 1998.

[241] Andrew Nealen, Matthias Müller, Richard Keiser, Eddy Boxerman, and Mark Carlson. Physically based deformable models in computer graphics. In *Computer graphics forum*, volume 25, pages 809–836. Wiley Online Library, 2006.

[242] Grand Roman Joldes, Adam Wittek, and Karol Miller. Real-time nonlinear finite element computations on gpu–application to neurosurgical simulation. *Computer methods in applied mechanics and engineering*, 199(49-52):3305–3314, 2010.

[243] Thomas Sangild Sørensen and Jesper Mosegaard. An introduction to gpu accelerated surgical simulation. In *International Symposium on Biomedical Simulation*, pages 93–104. Springer, 2006.

[244] Nvidia. JETSON TX2: Thermal Design Guide. https://developer.nvidia.com/embedded/dlc/jetson-tx2-thermal-design-guide, 2017. [Online; accessed 21-July-2019].

[245] Intel. Packaging Specifications and Dimensions Thermal Resistance: Cyclone V Devices. https://www.intel.com/content/www/us/en/programmable/support/literature/lit-index/lit-pkg/thermal.html?type=thermal&family=Cyclone_V, 2016. [Online; accessed 21-July-2019].

[246] Intel. Packaging Specifications and Dimensions Thermal Resistance: Stratix V Devices. https://toshiba.semicon-storage.com/info/docget.jsp?did=13417, 2016. [Online; accessed 21-July-2019].

[247] Toshiba. power mosfet thermal design and attachment of a thermal fin. https://toshiba.semicon-storage.com/info/docget.jsp?did=13417, 2018. [Online; accessed 21-July-2019].

[248] Tina Song. Li-ion 7700 mAh 12.6v SB10K10388 data sheet. https://static.lenovo.com/na/social-responsibility/Lenovo_MSDS_Finder/msds_type/SB10K10388_UN38.3.pdf, 2015. [Online; accessed 24-July-2019].

[249] Green Energy. 603450 Rechargable Lithium ION Technical Data Sheet. http://www.greenergy-battery.com/Specs/liion/603450.pdf, 2016. [Online; accessed 24-July-2019].